ON THE EXTREMALITY, UNIQUENESS AND OPTIMALITY OF TRANSFERENCE PLANS

STEFANO BIANCHINI AND LAURA CARAVENNA

Abstract. In this paper we consider the following standard problems appearing in optimal transportation theory:

- when a transference plan is extremal,
- when a transference plan is the unique transference plan concentrated on a set A,
- when a transference plan is optimal.

We show that these three problems can be studied with a general approach:

1. choose some necessary conditions, depending on the problem we are considering;
2. find a partition into sets B_α where these necessary conditions become also sufficient;
3. show that all the transference plans are concentrated on $\cup_\alpha B_\alpha$.

Explicit procedures are provided in the three cases above, the principal one being that the problem has an hidden structure of linear preorder with universally measurable graph.

As by sides results, we study the disintegration theorem w.r.t. family of equivalence relations, the construction of optimal potentials, a natural relation obtained from c-cyclical monotonicity.

Contents

1. Introduction 1
1.1. Explanation of the approach 3
1.2. Structure of the paper 4
2. Setting and general scheme 6
3. Extremality of transference plans 8
4. Uniqueness of transport plans 11
5. Optimality 19
5.1. Extension of the construction 23
6. Examples 23
6.1. Dependence w.r.t. the set Γ 23
6.2. Analysis of the transport problem in the quotient space 26
7. Existence of an optimal potential 30
Appendix A. Disintegration theorem in countably generated probability spaces 35
A.1. Characterization of the disintegration for a family of equivalence relations 38
Appendix B. Perturbation by cycles 39
B.1. Borel, analytic and universally measurable sets 39
B.2. General duality results 40
B.3. Decomposition of measures with 0 marginals 41
B.4. Examples 47
Appendix C. The c-cyclically monotone relation 48
Appendix D. Notation 54
References 55

1. Introduction

Let (X, Ω, μ), (Y, Σ, ν) be two countably generated probability spaces, and let $(X \times Y, \Omega \otimes \Sigma)$ be the product measurable space. Using standard results on measure space isomorphisms (see for example the

Date: October 28, 2009.

proof of the last theorem of [12]), in the following we assume that $(X,\Omega)=(Y,\Sigma)=([0,1],\mathcal{B})$, where $\mathcal{B}$ is the Borel σ-algebra.

Let $\mathcal{P}([0,1]^2)$ be the set of Borel probability measures on $[0,1]^2$, and let $\Pi(\mu,\nu)$ be the subset of $\mathcal{P}([0,1]^2)$ satisfying the marginal conditions $(P_1)_\sharp\pi=\mu$, $(P_2)_\sharp\pi=\nu$, where $P_1(x,y)=x$, $P_2(x,y)=y$ are the projection on X, Y:

$$\Pi(\mu,\nu):=\Big\{\pi\in\mathcal{P}([0,1]^2):(P_1)_\sharp\pi=\mu,(P_2)_\sharp\pi=\nu\Big\}.$$

For $\pi\in\Pi(\mu,\nu)$ we will denote by $\Gamma\subset[0,1]^2$ a set such that $\pi(\Gamma)=1$: as a consequence of the inner regularity of Borel measures, it can be taken σ-compact.

For any Borel probability measure π on $[0,1]^2$, let $\Theta_\pi\subset\mathbf{P}([0,1]^2)$ be the π-completion of the Borel σ-algebra. We denote with $\Theta(\mu,\nu)\subset\mathbf{P}([0,1]^2)$ the $\Pi(\mu,\nu)$-*universally measurable σ-algebra*: it is the intersection of all completed σ-algebras of the probability measures in $\Pi(\mu,\nu)$:

$$\Theta(\mu,\nu):=\bigcap\Big\{\Theta_\pi,\pi\in\Pi(\mu,\nu)\Big\}. \tag{1.1}$$

We define the functional $\mathcal{I}:\Pi(\mu,\nu)\to\mathbb{R}$ by

$$\mathcal{I}(\pi):=\int c(x,y)\pi(dxdy), \tag{1.2}$$

where $c:[0,1]^2\to[0,+\infty]$ is a $\Theta(\mu,\nu)$-measurable cost function. The set $\Pi^f(\mu,\nu)\subset\Pi(\mu,\nu)$ is the set of probability measures belonging to $\Pi(\mu,\nu)$ and satisfying the geometrical constraint $\mathcal{I}(\pi)<+\infty$.

The problems we are considering in the next sections are whether a given measure $\pi\in\Pi(\mu,\nu)$ satisfies one of the following properties:

- *it is extremal in $\Pi(\mu,\nu)$;*
- *it is the unique measure in $\Pi(\mu,\nu)$ concentrated on a given set $A\in\Theta((\mu,\nu))$;*
- *it is minimizing the functional $\mathcal{I}(\pi)$ in $\Pi(\mu,\nu)$.*

We can restrict our analysis to the set $\Pi^f(\mu,\nu)$, by

- defining $c(x,y)=\mathbb{I}_\Gamma$ for a particular set Γ with $\pi(\Gamma)=1$ in the first case,
- defining $c(x,y)=\mathbb{I}_A$ in the second case,
- assuming that $\mathcal{I}(\pi)<+\infty$ to avoid trivialities in the third case.

In all the above cases a necessary condition can be easily obtained, namely

- π is acyclic in the first case (Definition 3.2),
- π is A-acyclic in the second case (Definition 4.2),
- π is c-cyclically monotone in the third case (Definition 5.1).

Nevertheless, there are explicit examples showing that this condition is only necessary.

The kernel is the following idea (Lemma 2.5). Let $\pi\in\Pi(\mu,\nu)$ be a transference plan.

Theorem 1.1. *Assume that there are partitions $\{X_\alpha\}_{\alpha\in[0,1]}$, $\{Y_\beta\}_{\beta\in[0,1]}$ such that*

1. *for all $\pi'\in\Pi^f(\mu,\nu)$ it holds $\pi'(\cup_\alpha X_\alpha\times Y_\alpha)=1$,*
2. *the disintegration $\pi=\int\pi_\alpha m(d\alpha)$ of π w.r.t. the partition $\{X_\alpha\times Y_\alpha\}_{\alpha\in[0,1]}$ is strongly consistent,*
3. *in each equivalence class $X_\alpha\times Y_\alpha$ the measure π_α is extremal/unique/optimal in $\Pi(\mu_\alpha,\nu_\alpha)$, where*

$$\mu_\alpha:=(P_1)_\sharp\pi_\alpha,\quad \nu_\alpha:=(P_2)_\sharp\pi_\alpha.$$

Then π is extremal/unique/optimal.

The main tool is the Disintegration Theorem A.7 presented in Appendix A and applied to the partition $\{X_\alpha\times Y_\beta\}_{\alpha,\beta\in[0,1]}$. This partitions are constructed in order to satisfy Point (3).

Before explaining the meaning of the above conditions, we consider the following corollaries. Instead of partitions, we will equivalently speak of equivalence classes and relative equivalence relations.

Corollary 1.2 (Extremality (Theorem 3.8))**.** *Let π concentrated on a σ-compact acyclic set Γ.*

If we partition the set Γ into axial equivalence classes (Definition 3.4), then π is extremal in $\Pi(\mu,\nu)$ if the disintegration is strongly consistent.

We show in Theorem 3.9 that the strong consistency assumption in the above corollary is nothing more than the countable Borel limb condition of [11].

Denote with h_X, h_Y the quotient maps w.r.t. the partitions $\{X_\alpha\}_{\alpha\in[0,1]}$, $\{Y_\beta\}_{\beta\in[0,1]}$. In Lemma 2.4 it is shown that if Conditions (1) and (2) of Theorem 1.1 are valid for π, then there exists $m \in \mathcal{P}([0,1])$ such that $(\mathbb{I},\mathbb{I})_\sharp m = (h_X, h_X)_\sharp \mu = (h_Y, h_Y)_\sharp \nu = (h_X \otimes h_Y)_\sharp \pi$.

Let now A be an analytic set and define the image set

$$A' := (h_X \otimes h_Y)(A).$$

Corollary 1.3 (Uniqueness (Page 18)). *Let π concentrated on a σ-compact A-acyclic set Γ.*

If we partition the set Γ into axial equivalence classes, then π is the unique measure in $\Pi(\mu,\nu)$ concentrated on A if

(1) *the disintegration is strongly consistent,*
(2) *there exists a set $B \in (h_X \otimes h_Y)_\sharp \Theta(\mu,\nu)$, $B \supset A'$, which is the graph of a linear order.*

Notice that one can always take as a quotient space a subset of $[0,1]$, by the Axiom of Choice, but the image σ-algebra does not contain in general all Borel sets. Moreover, by Lemma 4.7, A' is by construction the graph of a partial order, which again by the Axiom of Choice can be always completed to a linear order. The two assumptions above are therefore a measurability assumption, made precise in Remark 4.19: A and Γ induce a preorder on $[0,1]$ which is contained in a linear (or total) preorder with Borel graph.

Finally, let $c : [0,1]^2 \to [0,+\infty]$ be a coanalytic cost.

Corollary 1.4 (Optimality (Theorem 5.6)). *Let π concentrated on a σ-compact c-cyclically monotone set Γ and partition Γ w.r.t. the cycle equivalence relation (Definition 5.1).*

Then, π c-cyclically monotone is optimal if

(1) *the disintegration is strongly consistent,*
(2) *the image set $A' := (h_X \otimes h_Y)(\{c < +\infty\})$ is a set of uniqueness.*

If we use as condition for uniqueness the linear order condition, then the interpretation of this corollary in terms of order relation is analogous to the one above and it is performed in Remark 5.10: $\{c < +\infty\}$ and Γ induce a preorder on $[0,1]$ which is contained in a linear preorder with Borel graph.

The above result generalizes the previous known cases:

(1) if μ or ν are atomic ([17]): clearly m must be atomic;
(2) if $c(x,y) \leq a(x) + b(y)$ with $a \in L^1(\mu)$, $b \in L^1(\nu)$ ([18]): m is a single δ;
(3) if $c : [0,1]^2 \to \mathbb{R}$ is real valued and satisfies the following assumption ([2])
$$\nu\bigg(\bigg\{y : \int c(x,y)\mu(dx) < +\infty\bigg\}\bigg) > 0, \quad \mu\bigg(\bigg\{x : \int c(x,y)\nu(dy) < +\infty\bigg\}\bigg) > 0 :$$
in this case m is a single δ;
(4) If $\{c < +\infty\}$ is an open set O minus a $\mu \otimes \nu$-negligible set N ([3]): in this case every point in $\{c < +\infty\}$ has a squared neighborhood of positive π-measure satisfying condition (5.4b) below.

In each case the equivalence classes are countably many Borel sets, so that the disintegration is strongly consistent and the acyclic set A' is a set of uniqueness (Lemma 4.18).

1.1. **Explanation of the approach.** The three conditions listed in Theorem 1.1 have interesting interpretations in terms of measurability, marginal conditions and acyclic perturbations.

We first observe that the *necessary conditions* considered in all three cases can be stated as follows: *the transference plan π is unique/optimal w.r.t. the affine space generated by $\pi + \lambda_c$, where λ_c is a cyclic perturbation of π.*

Moreover, *the partitions have a natural crosswise structure w.r.t.* Γ: if $\{X_\alpha\}_\alpha$, $\{Y_\beta\}_\beta$ are the corresponding decompositions of $[0,1]$, then

$$\Gamma \cap (X_\alpha \times Y) = \Gamma \cap (X \times Y_\alpha) = \Gamma \cap (X_\alpha \times Y_\alpha). \tag{1.3}$$

This is clearly equivalent to $\Gamma \subset \cup_\alpha X_\alpha \times Y_\alpha$, so that Condition (1) is satisfied at least for π and for its cyclic perturbations.

This and consequently Condition (1) are *conditions on the geometry of the carriage* Γ, since the specific construction depends on it. In fact, fixed a procedure to partition a set Γ, it is easy to remove negligible sets obtaining different partitions: sometimes Theorem 1.1 can be satisfied or not depending on Γ, i.e. on the partition. A possible solution is to make the partition independent of Γ (Appendix A.1), but maybe this decomposition does not satisfy the hypotheses of Theorem 1.1, while others do.

A consequence of the above discussion is that in the corollaries *a procedure is proposed to test a particular measure* π. Some particular cost may however imply that there is a partition valid for all transference plans: in this case the c-cyclical monotonicity becomes also sufficient, as in the known cases of Points (1)-(4) above.

Notice however that the statement is that *the necessary condition becomes sufficient if there exists a carriage* Γ *such that the corollaries apply*, or more generally if there exists a partition such that Theorem 1.1 applies. When there is no such carriage, then one can modify the cost in such a way that there are transport plans satisfying the necessary condition, giving the same quotient set A' and which can be either extremal/unique/optimal or not (Proposition 6.9).

The strong consistency of the disintegration is a measure theoretic assumption: it is equivalent to the fact that the quotient space can be taken to be $([0,1],\mathcal{B})$, up to negligible sets. This is important in order to give a meaning to the optimality within the equivalence classes: otherwise the conditional probabilities π_α are useless and Condition (3) without meaning. From the geometrical point of view, we are saying that π can be represented by weighted sum of probabilities in $X_\alpha \times Y_\alpha$, and Condition (1) yields that we can decompose the problem into smaller problems in $X_\alpha \times Y_\alpha$. When the assumption is not satisfied, then one can modify the cost in order to have the same quotient measure but both c-cyclically monotone optimal and c-cyclically non optimal transport plans (Example 6.5).

Finally, we illustrate *the linear preorder condition* of Corollary 1.3. The sets Γ and A (or $\{c < +\infty\}$) yield a natural preorder by saying $x \preccurlyeq x'$ if there exists an *axial path* connecting them:

$$\exists (x_i, y_i) \in \Gamma, i = 0, \dots, I \ : \ (x_{i+1}, y_i) \in A \ \forall i = 0, \dots, I \text{ and } x_0 = x, x_{I+1} = x'.$$

The equivalence classes $\{x \preccurlyeq x' \wedge x' \preccurlyeq x\}$ are the points connected by closed cycles. This holds in general, but a strong requirement is that this preorder can be embedded into a linear preorder (i.e. every two points are comparable) with Borel (universally measurable) graph having the same equivalence classes $\{x \preccurlyeq x' \wedge x' \preccurlyeq x\}$. If this holds, then Theorem 4.9 implies two things:

1. the disintegration with respect to the equivalence classes $\{x \preccurlyeq x' \wedge x' \preccurlyeq x\}$ is strongly consistent,
2. and the image set A' is contained in $B := \{s, t \in [0,1]^\alpha, s \trianglelefteq t\}$, with $\alpha \in \omega_1$ and $\trianglelefteq$ being the lexicographic ordering.

The last point and Lemma 4.13 (B is a set of uniqueness) prove that the assumption of Theorem 1.1 are verified.

1.2. **Structure of the paper.** The paper is organized as follows.

In Section 2 we show the general scheme of our approach. We do not specify the particular necessary conditions for optimality, but we prove that under the above three conditions the transference plan π is extremal/unique/optimal. In Section 2, page 8, we collect the results into 4 steps which will be used to obtain the results in the next sections.

In Section 3 we address the problem of extremality. The results obtained with our approach are already known in the literature: this part can be seen as an exercise to understand how the procedure works. The difficulties of both approaches are the same: in fact the existence of a Borel rooting set up to negligible sets is equivalent to the strong consistency of the disintegration.

In Section 4, we consider the problem of verifying if an analytic set A can carry more than one transference plan. In this case, not only the disintegration should be strongly consistent, but we must verify also Condition (1) of Theorem 1.1. Condition (2) in Corollary 1.3 implies this fact. Essentially, we are just showing (Theorem 4.9) that in the quotient space the uniqueness problem can be translated into the uniqueness problem in $[0,1]^\alpha$, with $\alpha \in \omega_1$ enumerable ordinal, and

$$B = \big\{(s,t) : s, t \in [0,1]^\alpha, s \trianglelefteq t\big\},$$

where $\trianglelefteq$ is the lexicographic order. Lemma 4.13 proves that the above B is a set of uniqueness.

In Section 5 we consider the optimality of a transference plan. In this case, the easiest equivalence relation is the cycle equivalence relation Definition (5.1), introduced also in [3]. The optimality within each class is immediate from the fact that there exists a couple of $\mathcal{A}$-optimal potentials ϕ,ψ, and after the discussion of the above two problems the statement of Corollary 1.4 should be clear. If one chooses the existence of optimal potentials as sufficient criterion, or even more general criteria, it is in general possible to construct other equivalence relations, such that in each class the conditional probabilities π_α are optimal. Under strong measure theoretic assumptions (ZFC+CH+PD), an example of this construction is shown in Appendix C.

In Section 6 we give several examples: for historical reasons, we restrict to examples concerning the optimality of π, but trivial variations can be done in order to adapt to the other two problems. We split the section into 2 parts. In Section 6.1 we study how the choice of Γ can affect our construction: it turns out that in pathological cases a wrong choice of Γ may lead to situations for which either the disintegration is not strongly consistent or in the quotient space there is no uniqueness. This may happen both for optimal or not optimal transference plans. In Section 6.2 instead we consider if one can obtain conditions on the problem in the quotient space less strict than the uniqueness condition: the examples show that this is not the case in general.

In Section 7 we address the natural question: if we have optimal potentials in each set $X_\alpha \times Y_\alpha$, is it possible to construct an optimal couple (ϕ, ψ) in $\cup_\alpha X_\alpha \times Y_\alpha$? We show that under the assumption of strong consistency this is the case. The main tool is Von Neumann's Selection Theorem, and the key point is to show that the set

$$\Big\{(\alpha, \phi_\alpha, \psi_\alpha) : \phi_\alpha, \psi_\alpha \text{ optimal couple in } X_\alpha \times Y_\alpha\Big\}$$

is analytic in a suitable Polish space. The Polish structure on the family of optimal couples is obtained identifying each μ-measurable function ϕ with the sequence of measures $\{(\phi \vee (-M)) \wedge M)\mu\}_{M\in\mathbb{N}}$, which is shown to be a Borel subset of $\mathcal{M}^{\mathbb{N}}$.

In Appendix A we give a short proof of the Disintegration Theorem in countably generated measure spaces. All the results of this section can be found in Section 452 of [10] (with much greater generality). In particular, the fact that consistent disintegrations exist and are unique, and the explicit representation of the conditional probabilities. As an application of these methods, we show that if one has a family of equivalence relations $\mathfrak{E}$ closed under countable intersection, then there is an equivalence relation $E \in \mathfrak{E}$ which is the sharpest one in the following sense: the σ-algebra of saturated sets w.r.t. any other $E' \in \mathfrak{E}$ can be embedded into the σ-algebra of saturated sets w.r.t. E (Point (1) Theorem A.11). Applied to our problem, we can make the disintegration independent of the particular carriage Γ, but the examples show that maybe this is not the best choice, or it is even the trivial one $x^\bullet = \{x\}$!

In Appendix B we give a meaning to the concepts of cyclic perturbations and acyclic perturbations. After recalling the properties of projective sets in Polish spaces in Section B.1 and the duality results of [13] (Section B.2), we show how to define the n-cyclic part of a signed measure λ with 0 marginals: this is the largest measure $\lambda_n \ll \lambda$ which can be written as $\lambda_n = \lambda_n^+ - \lambda_n^-$ with

$$\lambda_n^+ = \frac{1}{n} \int_{C_n} \sum_{i=1}^n \delta_{P_{(2i-1,2i)} w} m(dw) \qquad \lambda_n^- = \frac{1}{n} \int_{C_n} \sum_{i=1}^n \delta_{P_{(2i+1,2i \mod 2n)} w} m(dw)$$

where $C_n \subset [0,1]^{2n}$ is the set of n-closed cycles and $m \in \mathcal{M}^+(C_n)$. This approach leads to the definition of cyclic perturbations λ: these are the signed measures with 0 marginals which can be written as sum (without cancellation) of cyclic measures. The acyclic measures are those measures for which there are not n-cyclic measures $\lambda_n \ll \lambda$ for all $n \geq 2$: in particular they are concentrated on an acyclic set. This approach leads naturally to the well known results on the properties of sets on which extremal/unique/optimal measures are concentrated: in fact, in all cases we ask that there are not cyclic perturbations which either are concentrated on the carriage set Γ, or on the set of uniqueness A, or diminish the cost of the measure π. One then deduces the well known criteria that Γ is acyclic, Γ is A-acyclic and Γ is c-cyclically monotone.

The last Appendix C is more set theoretical: its aim is just to show that there are other possible decompositions for which our procedure can be applied, and in particular situations where a careful analysis may give the validity of Theorem 1.1 for this new decomposition, but not of Corollary 1.4 for the cycle decomposition. The main result is that under PD and CH we can construct a different equivalence relation satisfying Condition (3) of Theorem 1.1 and (1.3).

2. Setting and general scheme

Let $\{X_\alpha\}_{\alpha\in[0,1]}$ be a partition of X into pairwise disjoint sets, and similarly let $\{Y_\beta\}_{\beta\in[0,1]}$ be a partition of Y into pairwise disjoint sets. Let moreover $\{X_\alpha \times Y_\beta\}_{\alpha,\beta\in[0,1]}$ be the induced pairwise disjoint decomposition on $X \times Y$.

Since it is clear that the decomposition $X = \cup_\alpha X_\alpha$ with X_α pairwise disjoint induces an equivalence relation E by defining xEx' if and only if $x, x' \in X_\alpha$ for some α, we will also refer to X_α, Y_β and $X_\alpha \times Y_\beta$ as equivalence classes. We will often not distinguish an equivalence relation E on X and its graph

$$\mathrm{graph}(E) := \{(x,x') : xEx'\} \subset X \times X.$$

We will denote by $h_X : X \to [0,1]$, $h_Y : Y \to [0,1]$ the quotient maps: clearly $h_X \otimes h_Y : X \times Y \to [0,1]^2$ is the quotient map corresponding to the decomposition $X_\alpha \times Y_\beta$, $\alpha, \beta \in [0,1]$, of $X \times Y$.

Assumption 1. The maps h_X, h_Y are μ-measurable, ν-measurable from (X,Ω,μ), (Y,Σ,ν) to $([0,1],\mathcal{B})$, respectively, where $\mathcal{B}$ is the Borel σ-algebra.

We will consider the following disintegrations:

$$\mu = \int_0^1 \mu_\alpha m_X(d\alpha), \qquad m_X = (h_X)_\sharp \mu; \tag{2.1a}$$

$$\nu = \int_0^1 \nu_\beta m_Y(d\beta), \qquad m_Y = (h_Y)_\sharp \nu; \tag{2.1b}$$

$$\pi = \int_{[0,1]^2} \pi_{\alpha\beta} n(d\alpha d\beta), \qquad n = (h_X \otimes h_Y)_\sharp \pi. \tag{2.1c}$$

Note the fact that under the assumptions of measurability of h_X, h_Y, Theorem A.7 implies that — up to a redefinition of μ_α, ν_α, π_α on respectively m_X, m_Y, n negligible sets — the conditional probabilities μ_α, ν_β and $\pi_{\alpha,\beta}$ satisfy

$$\mu_\alpha(X_\alpha) = \nu_\beta(Y_\beta) = \pi_{\alpha\beta}(X_\alpha \times Y_\beta) = 1$$

for all $(\alpha,\beta) \in [0,1]^2$, i.e. they are concentrated on equivalence classes: in the following we will say that the disintegration is *strongly consistent* when the conditional probabilities are supported on the respective equivalence classes (see [10], Chapter 45, Definition 452E).

The next Lemma 2.1 is valid also in the case the disintegration is not strongly consistent but just consistent, by considering the quotient measure space of Definition A.5.

Lemma 2.1. *The measure n belongs to $\Pi(m_X, m_Y)$.*

Proof. This is a trivial consequence of the computation

$$n(A \times [0,1]) \overset{(2.1c)}{=} \pi(h_X^{-1}(A) \times Y) \overset{\pi\in\Pi(\mu,\nu)}{=} \mu(h_X^{-1}(A)) \overset{(2.1a)}{=} m_X(A).$$

The same computation works for $n([0,1] \times B)$. □

In the next sections, a special choice of the equivalence classes will lead to the following particular case, which under Assumption 1 is meaningful: indeed, as direct consequence of the properties of product σ-algebra (Theorem 3 in [12]), the set $\{\alpha = \beta\}$ belongs to the product σ-algebra $(h_X)_\sharp(\Omega) \otimes (h_Y)_\sharp(\Sigma)$ if and only if Assumption 1 holds (up to measure spaces isomorphisms).

Assumption 2. We assume $n = (\mathbb{I},\mathbb{I})_\sharp m_X$.
In particular the marginals m_X and m_Y coincide: we will denote this probability measure by m.

Hence the image of $\Pi(\mu,\nu)$ under $(h_X \otimes h_Y)$ is contained in the set $\Pi(m,m)$ by Lemma 2.1. Moreover:

Lemma 2.2. *Under Assumption 2, one has $\pi_\alpha \in \Pi(\mu_\alpha, \nu_\alpha)$.*

Proof. By the marginal conditions, for any m-measurable A and Borel S

$$\int_A \mu_\alpha(S) m(d\alpha) = \mu(h_X^{-1}(A) \cap S) \overset{\pi\in\Pi(\mu,\nu)}{=} \pi((h_X^{-1}(A) \cap S) \times [0,1]) = \int_A \pi_\alpha(S \times [0,1]) m(d\alpha).$$

Thus $(P_1)_\sharp \pi_\alpha = \mu_\alpha$ for m-a.e. α. For ν_α it is analogous. □

Under Assumption 1, a necessary and sufficient condition for Assumption 2 is the following.

Definition 2.3. We say that a set $\Gamma \subset [0,1]^2$ satisfies the *crosswise condition* w.r.t. the families $\{X_\alpha\}_{\alpha\in[0,1]}$, $\{Y_\beta\}_{\beta\in[0,1]}$, if

$$\Gamma \cap (X_\alpha \times Y) = \Gamma \cap (X \times Y_\alpha) = \Gamma \cap (X_\alpha \times Y_\alpha) \quad \forall \alpha \in [0,1]. \tag{2.2}$$

Lemma 2.4. *Assume that there exists $\Gamma \subset [0,1]^2$ such that $\pi(\Gamma) = 1$ and it satisfies the crosswise condition* (2.2)*. Then $n = (\mathbb{I},\mathbb{I})_\sharp m$, where $m = m_X = m_Y$.*

Conversely, if $n = (\mathbb{I},\mathbb{I})_\sharp m$, then there exists $\Gamma \subset [0,1]^2$ such that $\pi(\Gamma) = 1$ and satisfying (2.2)*.*

Proof. The proof follows the same line of the proof of Lemma 2.1.

The set $\Gamma' = (h_X \otimes h_Y)^{-1}(\{\alpha = \beta\})$ has full π measure if and only if $n = (h_X \otimes h_Y)_\sharp \pi = (\mathbb{I},\mathbb{I})_\sharp m$. Since (2.2) implies immediately $\Gamma \subset \Gamma'$, then $n = (\mathbb{I},\mathbb{I})_\sharp m$.

Conversely, by the definition of Γ'

$$(X_\alpha \times Y) \cap \Gamma' = \Gamma' \cap (X \times Y_\alpha) = X_\alpha \times Y_\alpha.$$

This implies (2.2) for the set Γ'. $\square$

Along with the strong consistency of the disintegration (Assumption 1), the main assumption is the following. This assumption requires Assumption 1 and implies Assumption 2.

Assumption 3. For all $\pi \in \Pi^f(\mu,\nu)$, the image measure $n = (h_X \otimes h_Y)_\sharp \pi$ is equal to $(\mathbb{I},\mathbb{I})_\sharp m$.

So far we do not have specified the criteria to choose the partitions X_α, Y_β. The next lemma, which is the key point of the argument, specifies it.

Lemma 2.5. *Assume that the decompositions X_α, Y_β satisfy Assumption 3 and the following:*

Assumption 4. For m-a.e $\alpha \in [0,1]$ the probability measure $\pi_\alpha \in \Pi(\mu_\alpha,\nu_\alpha)$ satisfies sufficient conditions for extremality/uniqueness/optimality.

Then $\pi \in \Pi(\mu,\nu)$ is extremal/unique/optimal.

Proof. We consider the cases separately.

Extremality. If $\pi_1, \pi_2 \in \Pi(\mu,\nu)$ are such that $\pi = (1-\lambda)\pi_1 + \lambda\pi_2$, $\lambda \in (0,1)$, then it follows from Assumption 3 that the disintegration of these measures is given by

$$\pi_1 = \int_0^1 \pi_{1,\alpha} m(d\alpha), \quad \pi_2 = \int_0^1 \pi_{2,\alpha} m(d\alpha) \qquad \pi_{1,\alpha}, \pi_{2,\alpha} \in \Pi(\mu_\alpha,\nu_\alpha) \text{ by Lemma 2.2.}$$

It follows that $\pi_\alpha = (1-\lambda)\pi_{1,\alpha} + \lambda\pi_{2,\alpha}$ for m-a.e. α, so that from Assumption 4 we conclude that $\pi_\alpha = \pi_{1,\alpha} = \pi_{2,\alpha}$.

Uniqueness. The computations are similar to the previous case, only using the fact that in each class the conditional probability π_α is unique.

Optimality. For $\pi_1 \in \Pi^f(\mu,\nu)$

$$\mathcal{I}(\pi_1) = \int c(x,y)\pi_1(dxdy) \overset{(2.1c)}{=} \int_0^1 \bigg(\int c(x,y)\pi_{1,\alpha}(dxdy) \bigg) m(d\alpha).$$

From Assumption 3 it follows that $\pi_{1,\alpha}, \pi_\alpha \in \Pi(\mu_\alpha,\nu_\alpha)$, so that from Assumption 4 one has

$$\int c(x,y)\pi_{1,\alpha}(dxdy) \geq \int c(x,y)\pi_\alpha(dxdy) \qquad \text{for } m\text{-a.e. } \alpha.$$

The conclusion follows. $\square$

We thus are left to perform the following steps in each of the next sections.

Procedure to verify the sufficiency of the necessary conditions.

(1) Fix the necessary conditions under consideration.
(2) Fix a measure $\pi \in \Pi^f(\mu,\nu)$ which satisfies the necessary conditions respectively for being extremal, being the unique measure concentrated on A, being optimal.
(3) Construct partitions X_α, Y_β of X, Y such that:
   (a) the disintegrations of μ, ν w.r.t. $X = \cup_\alpha X_\alpha$, $Y = \cup_\beta Y_\beta$ are strongly consistent. This implies that the quotient maps h_X, h_Y can be assumed to be measurable functions taking values in $([0,1],\mathcal{B})$, by Theorem A.7;
   (b) in each equivalence class $X_\alpha \times Y_\alpha$ the necessary conditions become sufficient: the measure $\pi_{\alpha\alpha}$ satisfies the sufficient conditions for extremality, uniqueness or optimality among all $\pi \in \Pi(\mu_\alpha,\nu_\alpha)$.
(4) Verify that the image measure $n_{\pi'} \in \Pi(m,m)$ of all $\pi' \in \Pi^f(\mu,\nu)$ coincides with $(\mathbb{I},\mathbb{I})_\sharp m$, where $m = (h_X)_\sharp\mu = (h_Y)_\sharp\nu$.

If the above steps can be performed, then from Lemma 2.5 we deduce that π is respectively extremal, unique or optimal. In our applications, the necessary conditions reduce to a single condition on the structure of the support of π.

Remark 2.6. It is important to note that in general the decomposition depends on the particular measure π under consideration: the procedure will be used to test a particular measure π, even if in some cases it works for the whole $\Pi^f(\mu,\nu)$. In the latter case, we can test e.g. the optimality of all measures in $\Pi^f(\mu,\nu)$ using only the necessary conditions: this means that these conditions are also sufficient.

3. Extremality of transference plans

The first problem we will consider is to give sufficient conditions for the extremality of transference plans in $\Pi(\mu,\nu)$. The results obtained are essentially the same as the results of [11].

We first recall the following result ([7, 14]). Following the notation of Appendix B.3, we denote with $\Lambda \subset \mathcal{M}([0,1]^2)$ the set

$$\Lambda := \Big\{\lambda \in \mathcal{M}([0,1]^2) : (P_1)_\sharp\lambda = (P_2)_\sharp\lambda = 0\Big\}.$$

Proposition 3.1. *The transference plan $\pi \in \Pi(\mu,\nu)$ is extremal if and only if $L^1(\mu) + L^1(\nu)$ is dense in $L^1(\pi)$.*

Proof. We first prove that if $f_1 \in L^1(\mu)$, $f_2 \in L^1(\nu)$ and $(f_1 - f_2)\pi \in \Lambda$, then $f_1 - f_2 = 0$ π-a.e..

Writing

$$\pi = \int \pi_x \mu(dx) = \int \pi_y \nu(dy)$$

for the disintegration of π w.r.t. μ, ν respectively, the above conditions mean that

$$f_1(x) = \int f_2(y)\pi_x(dy) \quad \mu\text{-a.e. } x, \qquad f_2(y) = \int f_1(x)\pi_y(dx) \quad \nu\text{-a.e. } y.$$

We then have

$$\begin{aligned}\int |f_1|\mu &= \int \bigg| \int f_2(y)\pi_x(dy) \bigg| \mu(dx) \\ &= \int |f_2|\nu + \int \bigg(\bigg| \int f_2(y)\pi_x(dy) \bigg| - \int |f_2(y)|\pi_x(dy) \bigg) \mu(dx) \leq \int |f_2|\nu,\end{aligned}$$

and similarly

$$\int |f_2|\nu = \int |f_1|\mu + \int \bigg(\bigg| \int f_1(x)\pi_y(dx) \bigg| - \int |f_1(x)|\pi_y(dx) \bigg) \nu(dy) \leq \int |f_1|\mu.$$

We thus conclude that

$$\bigg| \int f_2(y)\pi_x(dy) \bigg| = \int |f_2(y)|\pi_x(dy) \quad \mu \text{ a.e. } x, \qquad \bigg| \int f_1(x)\pi_y(dx) \bigg| = \int |f_1(x)|\pi_y(dx) \quad \nu \text{ a.e. } y.$$

i.e. π is concentrated on the set

$$\{f_1 < 0\} \times \{f_2 < 0\} \cup \{f_1 = 0\} \times \{f_2 = 0\} \cup \{f_1 > 0\} \times \{f_2 > 0\}.$$

Since if (f_1, f_2) satisfies $(f_1 - f_2)\pi \in \Lambda$, also $[(f_1 - k) - (f_2 - k)]\pi \in \Lambda$ for all $k \in \mathbb{R}$, it follows that π is concentrated on the sets

$$\{f_1 < k\} \times \{f_2 < k\} \cup \{f_1 = k\} \times \{f_2 = k\} \cup \{f_1 > k\} \times \{f_2 > k\}.$$

Hence one concludes that $f_1 - f_2 = 0$ π a.e..

$\Longleftarrow$ The previous step implies that if $L^1(\mu) + L^1(\nu)$ is dense in $L^1(\pi)$, then π should be extremal. In fact, it is fairly easy to see that if π is not extremal, then there exists $0 \leq g \in L^1(\pi)$ such that $g\pi \in \Pi(\mu, \nu)$: hence for some sequence $\{(f_{1,n}, f_{2,n})\}_{n\in\mathbb{N}} \in L^1(\mu) \times L^1(\nu)$ it holds

$$\lim_{n\to\infty} \int |g - f_{1,n} - f_{2,n}|\pi = 0. \tag{3.1}$$

Define the $L^1(\mu)$-function $m_{1,n}$ and the $L^1(\nu)$-function $m_{2,n}$ by

$$m_{1,n}\mu := (P_1)_\sharp\big(g - f_{1,n} - f_{2,n}\big)\pi, \quad m_{2,n}\mu := (P_2)_\sharp\big(g - f_{1,n} - f_{2,n}\big)\pi.$$

From (3.1), it follows that

$$\lim_{n\to+\infty} \|m_{1,n}\|_{L^1(\mu)} + \|m_{2,n}\|_{L^1(\nu)} = 0. \tag{3.2}$$

Trivially we have

$$f_{1,n} + m_{1,n} + f_{2,n} + m_{2,n} \in L^1(\pi), \quad \big(f_{1,n} + m_{1,n} + f_{2,n} + m_{2,n}\big)\pi \in \Pi(\mu, \nu).$$

Hence by the first part up to a π-negligible set

$$f_{1,n} + f_{2,n} = 1 - (m_{1,n} + m_{2,n}).$$

From (3.2) it follows that $f_{1,n} + f_{2,n} \to 1 = g$ in $L^1(\pi)$.

$\Longrightarrow$ If instead $\overline{L^1(\mu) + L^1(\nu)} \subsetneq L^1(\pi)$, then by Hahn-Banach Theorem there exists an $L^\infty(\pi)$-function $g \neq 0$, $|g| \leq 1$, such that

$$\int g(x,y)(f_1(x) + f_2(y))\pi(dxdy) = 0$$

for all $f_1 \in L^1(\mu)$, $f_2 \in L^1(\nu)$. In particular $g\pi \in \Lambda$, $g \neq 0$ on a set of positive π-measure and

$$\pi = \frac{1+g}{2}\pi + \frac{1-g}{2}\pi,$$

where the two addends in the r.h.s. above belongs to $\Pi(\mu/2, \nu/2)$. $\square$

The second result is a consequence of Proposition B.15. A cyclic perturbation λ of a measure $\pi \in \Pi(\mu, \nu)$ is specified in Definitions B.6, B.14; in particular $\pi + \lambda \in \Pi(\mu, \nu)$.

Definition 3.2 (Acyclic set and measure)**.** We say that $\Gamma \subset [0,1]^2$ is *acyclic* if for all finite sequences $(x_i, y_i) \in \Gamma$, $i = 1, \dots, n$, with $x_i \neq x_{i+1 \mod n}$ and $y_i \neq y_{i+1 \mod n}$ it holds

$$\Big\{(x_{i+1}, y_i), i = 1, \dots, n, x_{n+1} = x_1\Big\} \not\subset \Gamma.$$

A measure is *acyclic* if it is concentrated on an acyclic set.

Lemma 3.3 (Theorem 3 of [11])**.** *Suppose that there is no cyclic perturbation of the measure $\pi \in \Pi(\mu, \nu)$ on $[0,1]^2$. Then π is concentrated on an acyclic σ-compact set Γ.*

We specify now necessary and sufficient conditions for extremality:

- **necessary condition:** the measure π is acyclic;
- **sufficient condition:** the measure π is concentrated on a *Borel limb numbering system*, [11] page 223: there are two disjoint families $\{C_k\}_{k\in\mathbb{N}}$, $\{D_k\}_{k\in\mathbb{N}_0}$ of Borel sets and Borel measurable functions $f_k : C_k \to D_{k-1}$, $g_k : D_k \to C_k$, $k \in \mathbb{N}$, such that π is concentrated on the union of the following graphs
$$F_k = \mathrm{graph}(f_k), \quad G_k = \mathrm{graph}(g_k).$$

We verify directly the second condition, [11] Theorem 20: clearly due to the σ-additivity and inner regularity, we can always replace measurable with σ-compact sets up to a negligible set.

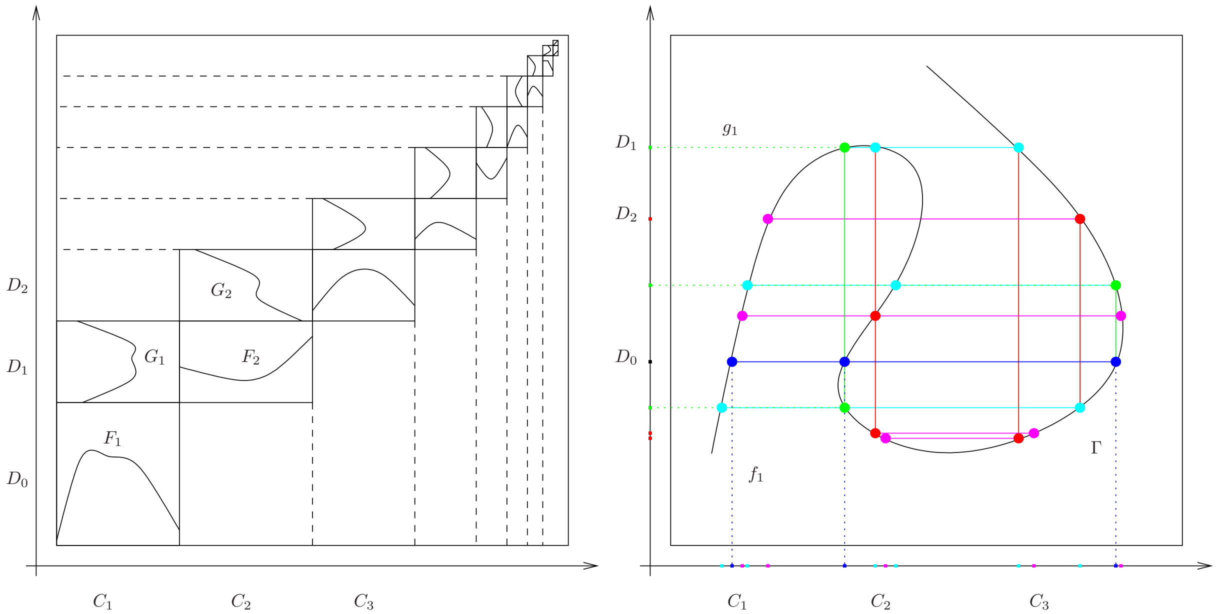


Figure 1: A limb numbering system and the axial path of a point.

Proof of sufficiency of the condition. Assume first that there are only finitely many G_k, F_k, $k \leq N$. In this case, the uniqueness of the transference plan π follows by finite recursion, since the marginality conditions yield, setting $F_{N+1} := \emptyset$, that π must be defined by

$$\pi \llcorner_{F_k} = (\mathbb{I}, f_k)_\sharp(\mu - (P_1)_\sharp \pi \llcorner_{G_k}), \quad \pi \llcorner_{G_k} = (g_k, \mathbb{I})_\sharp(\nu - (P_2)_\sharp \pi \llcorner_{F_{k+1}}), \qquad k \in \{1, \dots, N\}. \tag{3.3}$$

For the general case, let $\pi \in \Pi(\mu, \nu)$ such that $\pi(\cup_k F_k \cup G_k) = 1$. Define the measures π_N by means of (3.3) starting at N: let

$$(\pi_N)\llcorner_{F_{N+1}} := (\mathbb{I}, f_{N+1})_\sharp \mu \llcorner_{F_{N+1}}$$

and for $k \in \{1, \dots, N\}$

$$(\pi_N)\llcorner_{F_k} := (\mathbb{I}, f_k)_\sharp(\mu - (P_1)_\sharp(\pi_N)\llcorner_{G_k}), \quad (\pi_N)\llcorner_{G_k} = (g_k, \mathbb{I})_\sharp(\nu - (P_2)_\sharp(\pi_N)\llcorner_{F_{k+1}}).$$

Since $\sum_{k>N} \mu(F_k) + \nu(G_k) \to 0$ as $N \to \infty$, it is fairly easy to see that up to subsequences π_N converges strongly to π — just by the fact that if $\sum_i a_i < +\infty$ there exists a subsequence $i(j)$ s.t. $i(j)a_{i(j)} \to 0$.

Using the uniqueness of the limit and the fact that the approximating sequence does not depend on π, the uniqueness of π follows. □

The equivalence classes in order to apply Theorem 1.1 are the following.

Definition 3.4 (Axial equivalence relation)**.** We define $(x, y)E(x', y')$ if there are $(x_i, y_i) \in \Gamma$, $0 \leq i \leq I$ finite, such that

$$(x, y) = (x_0, y_0), (x', y') = (x_I, y_I) \quad \text{and} \quad (x_{i+1} - x_i)(y_{i+1} - y_i) = 0. \tag{3.4}$$

In the language of [11], page 222, each equivalence class is an *axial path*. The next lemma is an elementary consequence of Definition 3.4.

Lemma 3.5. *The relation E of Definition 3.4 defines an equivalence relation on the acyclic set Γ. If $\Gamma = \cup_\alpha \Gamma_\alpha$ is the partition of Γ in equivalence classes, and $X_\alpha = P_1 \Gamma_\alpha$, $Y_\alpha = P_2 \Gamma_\alpha$ are the projections of the equivalence classes, then the crosswise condition* (2.2) *holds.*

By setting

$$X_0 = [0, 1] \setminus P_1(\Gamma), \quad Y_0 = [0, 1] \setminus P_2(\Gamma),$$

we have a partition of X, Y into disjoint classes.

We can thus use Theorem A.7 to disintegrate the marginals μ, ν and every transference π plan supported on Γ. From (2.1) and Lemmas 2.2, 2.4 one has immediately the following proposition.

Proposition 3.6. *The following disintegrations w.r.t. the partitions $X = \cup_\alpha X_\alpha$, $Y = \cup_\alpha Y_\alpha$ hold:*

$$\mu = \int \mu_\alpha m(d\alpha), \quad \nu = \int \nu_\alpha m(d\alpha), \quad m = (h_X)_\sharp \mu.$$

Moreover, if π is a transference plan supported on Γ and the disintegration is strongly consistent, then the disintegration of π w.r.t. the partition $\Gamma = \cup_{\alpha \in A} \Gamma_\alpha$ is given by

$$\pi = \int \pi_\alpha m(d\alpha) \quad \text{with } \pi_\alpha \in \Pi(\mu_\alpha, \nu_\alpha).$$

The next lemma shows that in each equivalence class the sufficient condition holds.

Lemma 3.7. *Each equivalence class satisfies the Borel limb numbering condition.*

Proof. The proof is elementary: if $(x_\alpha, y_\alpha) \in \Gamma_\alpha$, then one defines recursively (Figure 1)

$$D_{0,\alpha} = \{y_\alpha\}, \quad C_{1,\alpha} = P_1(\Gamma \cap ([0,1] \times \{y_\alpha\})),$$

$$D_{k,\alpha} = P_2\big(\Gamma \cap (C_{k,\alpha} \times ([0,1] \setminus D_{k-1,\alpha}))\big), \quad C_{k+1,\alpha} = \Gamma \cap (([0,1] \setminus C_{k,\alpha}) \times D_{k,\alpha}).$$

From the assumption of acyclicity, it is a straightforward verification that each

$$\Gamma \cap (C_{k,\alpha} \times D_{k-1,\alpha}), \qquad \Gamma \cap (C_{k,\alpha} \times D_{k,\alpha})$$

is the graph of a function $f_{k,\alpha} : C_{k,\alpha} \to D_{k-1,\alpha}$, $g_{k,\alpha} : D_{k,\alpha} \to C_{k,\alpha}$. Moreover, Γ_α is covered by the graphs $G_{k,\alpha}$, $F_{k,\alpha}$ because of the definition of the equivalence class Γ_α.

It remains to study the Borel measurability of the functions $g_{k,\alpha}$, $f_{k,\alpha}$. We show the induction step of the argument. $D_{0,\alpha}$ is a point and $C_{1,\alpha}$ a section of the Borel set Γ, thus is itself Borel. Assume the $C_{k,\alpha}$, $D_{k-1,\alpha}$ are Borel. Then $\Gamma \cap (C_{k,\alpha} \times ([0,1] \setminus D_{k-1,\alpha})$ is the Borel antigraph $G_{k,\alpha}$: hence its horizontal section is compact, being a point, and by Novikov Theorem 4.7.11 its projection $D_{k,\alpha}$ is Borel. Finally by Theorem 4.5.2 of [19] the function $g_{k,\alpha}$ is Borel. The argument for $C_{k+1,\alpha}$ is analogous with $F_{k+1,\alpha}$. □

From Lemma 2.5, it follows the following theorem.

Theorem 3.8. *If the disintegration of Proposition 3.6 is strongly consistent, then π is extremal.*

We now conclude the section showing that the existence of a Borel limb numbering systems is equivalent to the existence of an acyclic set Γ where the transference plan π is concentrated and such that the disintegration is consistent.

Theorem 3.9. *The transference plans π is concentrated on a limb numbering system Γ with Borel limbs if and only if the disintegration of π into the equivalence classes of some acyclic carriage Γ is strongly consistent.*

Proof. Assume first that π satisfies the Borel limb condition. Then from [11], Theorem 20, it follows we can take as quotient space a Borel root set $\mathtt{A}$. In particular Γ can be taken as the union of the orbits of points in $\mathtt{A}$, and it is immediate to verify that the orbit of a Borel subset of $\mathtt{A}$ is an analytic subset of $[0,1]^2$. Hence the disintegration is consistent by the fact that $(\mathtt{A}, \mathcal{B}(A), m)$ is a countably generated measure space.

Conversely, suppose that the disintegration is strongly consistent w.r.t. the axial equivalence relation E on an acyclic carriage Γ. Then, as a consequence of Proposition A.9, by eventually removing a set of π-measure 0 from Γ, one can assume that the equivalence relation E has a Borel section S. One constructs finally Borel limbs as in Lemma 3.7 from $\{(x(\alpha), y(\alpha)\}_{\alpha \in [0,1]}$. □

Remark 3.10. We observe that by adding the set $G_0 = \{x_0\} \times D_0$, where $x_0 \notin \cup_k C_k$, the disintegration is supported on a single equivalence class.

4. Uniqueness of transport plans

In this section we address the question of uniqueness of transference plans concentrated on a set A.

Definition 4.1 (Set of uniqueness)**.** We say that $A \in \Theta(\mu, \nu)$ is a *set of uniqueness* of $\Pi(\mu, \nu)$ if there exists a unique measure $\pi \in \Pi(\mu, \nu)$ such that $\pi(A) = 1$.

In Section 5 of [11] (or using directly the proof of the sufficient condition, page 10) it is shown that if Γ satisfies the Borel limb condition, then Γ supports a unique transference plan.

The first lemma is a consequence of Proposition B.15.

Definition 4.2. A set $\Gamma \subset A$ is *A-acyclic* if for all finite sequences $(x_i, y_i) \in \Gamma$, $i = 1, \dots, n$, with $x_i \neq x_{i+1 \mod n}$ and $y_i \neq y_{i+1 \mod n}$ it holds

$$\Big\{ (x_{i+1}, y_i), i = 1, \dots, n, x_{n+1} = x_1 \Big\} \not\subset A.$$

A measure is *A-acyclic* if it is concentrated on an A-acyclic set.

Lemma 4.3. *If an analytic set A is a set of uniqueness for $\Pi(\mu, \nu)$, then the unique $\pi \in \Pi(\mu, \nu)$ is concentrated on a A-acyclic Borel set $\Gamma \subset A$.*

Necessary and sufficient conditions for uniqueness are then given by:

- **necessary condition:** there exist a measure $\pi \in \Pi(\mu, \nu)$ and an A-acyclic Borel set $\Gamma \subset A$ such that $\pi(\Gamma) = 1$;
- **sufficient condition:** A is a Borel limb numbering system (Page 9).

We will state a more general sufficient condition later at Page 18.

Let Γ be an A-acyclic σ-compact carriage of π. In particular, Γ is acyclic. We will thus use the equivalence classes of the axial equivalence relation E on Γ, Definition 3.4, assuming w.l.o.g. that $P_X(\Gamma) = P_Y(\Gamma) = [0, 1]$.

Let $h_X : X \to [0, 1]$, $h_Y : Y \to [0, 1]$ be the quotient maps. In general the image of A

$$A' := \Big\{ (\alpha, \beta) : (h_X \otimes h_Y)^{-1}(\alpha, \beta) \cap A \neq \emptyset \Big\} \tag{4.1}$$

is not a subset of $\{\alpha = \beta\}$. However, for the equivalence classes in the diagonal $\{\alpha = \beta\}$, we have the following lemma.

Lemma 4.4. *For all $\alpha \in [0, 1]$,*

$$(h_X \otimes h_Y)^{-1}(\alpha, \alpha) \cap A = (h_X \otimes h_Y)^{-1}(\alpha, \alpha) \cap \Gamma.$$

Proof. The definition implies that if $x, x' \in h_X^{-1}(\alpha)$, then there exist $(x_i, y_i) \in \Gamma$, $i = 0, \dots, I$, with $x_0 = x$, such that denoting $x_I = x'$ then (3.4) holds. A completely similar condition is valid for $y, y' \in h_Y^{-1}(\alpha)$.

Let $(\bar{x}, \bar{y}) \in (h_X^{-1}(\alpha) \times h_Y^{-1}(\alpha)) \cap (A \setminus \Gamma)$. Then there are $(x, y), (x', y') \in \Gamma$ such that $x = \bar{x}$, $y' = \bar{y}$. Consider then the axial path $(x_i, y_i) \in \Gamma$, $i = 0, \dots, I = 2(n-1)$, connecting them inside the class α: removing by chance some points, we can assume that $(x_0, y_0) = (x, y)$, $(x_I, y_I) = (x', y')$ and

$$x_{2j} - x_{2j-1} = 0, \quad y_{2j-1} - y_{2j-2} = 0, \quad j = 1, \dots, n.$$

Hence if we add the point $(x_{I+1}, y_{I+1}) = (\bar{x}, \bar{y})$ we obtain a closed cycle, contradicting the hypotheses of acyclicity of Γ in A. $\square$

The above lemma together with Lemma 2.5 and Lemma 3.7 implies that non uniqueness occurs because of the following two reasons:

(1) either the disintegration is not strongly consistent,
(2) or the push forward of some transference plan $\pi \in \Pi(\mu, \nu)$ such that $\pi(A) = 1$ is not supported on the diagonal in the quotient space.

Indeed, differently from the previous section, *the consistency of the disintegration is not sufficient to deduce the uniqueness of the transference plan.*

Example 4.5 (Pratelli). Consider $\mu = \mathcal{L}^1$ and the set

$$A = \{x = y\} \cup \{y - x = \alpha \mod 1\}, \quad \Gamma = \{x = y\} \quad \text{with } \alpha \in [0, 1] \setminus \mathbb{Q}.$$

In this case the quotient map is the identity, but the measure $(x, x + \alpha \mod 1)_\sharp \mathcal{L}^1$ is not concentrated on the diagonal and still belongs to $\Pi(\mathcal{L}^1, \mathcal{L}^1)$.

In the following we address the second point, and we assume that the disintegration is strongly consistent — which is equivalent to assume that the quotient maps h_X, h_Y can be taken Borel (up to a μ, ν negligible set, respectively, consequence of Proposition A.9).

Lemma 4.6. *The set A' defined in* (4.1) *is analytic if A is analytic.*

Proof. Since $A' = (h_X, h_Y)(A)$, the proof is a direct consequence of the fact that Borel images of analytic sets are analytic, being the projection of a Borel set. □

The next lemma is a consequence of the acyclicity of Γ in A.

Lemma 4.7. *In the quotient space, the diagonal is A'-acyclic.*

Proof. We prove the result only for 2-cycles, the proof being the same for the n-cycles.

Assume that A' has a 2-cycle, between the classes (α, α) and (α', α'). This means that there are points $(x, y) \in (h_X \otimes h_Y)^{-1}(\alpha, \alpha') \cap A$ and $(x', y') \in (h_X \otimes h_Y)^{-1}(\alpha', \alpha) \cap A$.

By definition of equivalence class, there are points $(x_i, y_i) \in (h_X \otimes h_Y)^{-1}(\alpha, \alpha)$, $i = 1, \dots, n$, and $(x'_j, y'_j) \in (h_X \otimes h_Y)^{-1}(\alpha', \alpha')$, $j = 1, \dots, n'$ forming an axial path in Γ and connecting (x, y) to (x', y') in $(h_X \otimes h_Y)^{-1}(\alpha, \alpha)$ and (x, y) to (x', y') in $(h_X \otimes h_Y)^{-1}(\alpha', \alpha')$.

The composition of the two axial paths yields a closed cycle, contradicting the assumption of acyclicity of Γ in A. □

We now give a sufficient condition for the implication

$$n \in \Pi(m, m), \ n(A') = 1 \quad \Longrightarrow \quad n(\{\alpha = \beta\}) = 1,$$

where $m = (h_X)_\sharp \mu = (h_Y)_\sharp \nu$.

Definition 4.8. A relation $R \subset [0, 1]^2$ is a *preorder* if

$$(x, y), (y, z) \in R \ \Rightarrow \ (x, z) \in R.$$

A preorder R is a *linear preorder* if $R \cup R^{-1} = [0, 1]^2$.

When the preorder is linear we are thus requiring that every couple is comparable. This means that R is a linear order when $[0, 1]$ is quotiented w.r.t. the equivalence relation

$$E := R \cap R^{-1}. \tag{4.2}$$

We will also write $x \preccurlyeq y$ or xRy when $(x, y) \in R$ and R is a preorder.

Notice that since A' is acyclic w.r.t. the diagonal (Lemma 4.7), it defines a partial order on $[0, 1]$ with analytic graph (Lemma 4.6).

Theorem 4.9. *Let B be the Borel graph of a linear preorder $\preccurlyeq$ on $[0, 1]$. Then the disintegration w.r.t. E is strongly consistent for all $\mu \in \mathcal{P}([0, 1])$, and the image set B' in the quotient space is a set of uniqueness of $\Pi(m, m)$, where m is the image measure of μ w.r.t. the equivalence relation E in* (4.2)*.*

In the proof, at Page 16, we use the following lemmas.

Lemma 4.10. *Let $m \in \mathcal{P}([0, 1])$ and $B \in \mathcal{B}([0, 1]^2)$. Then the function $x \mapsto h_B(x) := m(B(x))$ is Borel.*

Proof. First observe that if

$$B = \bigcup_{i=1}^{n} A_i \times A'_i, \quad A_i \times A'_i \cap A_j \times A'_j = \emptyset \text{ for } i \neq j, \ A_i, A'_i \in \mathcal{B}([0, 1]) \ \forall i \in 1, \dots, n.$$

then h_B is Borel:

$$h_B(x) = \sum_{i=1}^{n} m(A'_i) \chi_{A_i}(x).$$

Hence h_B is Borel on the algebra of simple products.

Moreover, if $\{B_n\}_{n \in \mathbb{N}}$ is an increasing sequence of Borel sets, then by the σ-additivity of the measure

$$h_B(x) = \sup_n h_{B_n}(x).$$

The same computation holds for a decreasing family of Borel sets $\{B_n\}_{n \in \mathbb{N}}$, $h_B(x) = \inf_n h_{B_n}(x)$.

It thus follows that the family of sets $\mathscr{A}$ such that h_B is Borel contains the simple products and it is a monotone class. From the Monotone Class Theorem (Proposition 3.1.14, page 85 of [19]) it follows that $\mathscr{A} \supset \mathcal{B}([0, 1]^2)$. □

Clearly, if $B \subset [0,1]^2$ is $\mu \otimes m$-measurable then by Fubini Theorem the function $x \mapsto h_B(x)$ is μ-measurable.

Lemma 4.11. *If $[0,1] \ni t \mapsto m_t \in \mathcal{P}([0,1])$ is Borel, then for all $B \in \mathcal{B}([0,1])$ the function $t \mapsto m_t(B)$ is Borel.*

Proof. Let $\mathscr{A} \subset \mathbb{P}([0,1])$ be the family of sets such that $t \mapsto m_t(B)$ is Borel for $B \in \mathscr{A}$.

If O is open, then the function $t \mapsto m_t(O)$ is l.s.c., being the supremum of continuous functions

$$m_t(O) = \sup_{\phi \in C([0,1])} \bigg\{ \int \phi m_t, \phi \leq \chi_O \bigg\},$$

so that open sets belong to $\mathscr{A}$.

Using the equivalences

$$m_t\bigg(\bigcup_{i\in\mathbb{N}} B_i\bigg) = \lim_{i\to\infty} m_t(B_i) \;\; \{B_i\}_i \text{ increasing}, \quad m_t([0,1]\setminus B) = 1 - m_t(B),$$

it follows that $\mathscr{A}$ is a σ-algebra. □

Lemma 4.12. *Let $[0,1] \ni t \mapsto m_t \in \mathcal{P}([0,1])$ be Borel and $B \in \mathcal{B}([0,1]^2)$. Then the function*

$$\begin{array}{cccc} h & : & [0,1]^2 & \to & [0,1] \\ & & (t,x) & \mapsto & h(t,x) := m_t(B(x)) \end{array}$$

is Borel.

Proof. If B is a finite union of disjoint products of Borel sets $A_i \times A_i'$, $i = 1, \dots, n$, then

$$h_B(t,x) = \sum_{i=1}^n m_t(A_i')\chi_{A_i}(x),$$

so that by Lemma 4.11 the function h_B is Borel on the algebra of simple products. As in Lemma 4.10, the family $\mathscr{A}$ of sets for which $h_B(t,x) = m_t(B(x))$ is Borel is a monotone class, so that the conclusion follows. □

In the following we will consider the set $[0,1]^\alpha$, where α is an ordinal number. The linear order on this set is the lexicographic ordering $\trianglelefteq$:

$$s,t \in [0,1]^\alpha,\; s \trianglelefteq t,\; s \neq t \quad \Longleftrightarrow \quad \exists \beta \leq \alpha\Big(\big(\forall \gamma < \beta(P_\gamma(s) = P_\gamma(t))\big) \wedge \big(P_\beta(s) < P_\beta(t)\big)\Big). \tag{4.3}$$

We recall that $P_\gamma : [0,1]^\alpha \to [0,1]$ is the projection on the γ-coordinate.

Lemma 4.13. *If $\alpha \in \omega_1$, then*

$$B_\alpha := \big\{(s,t) \in [0,1]^\alpha \times [0,1]^\alpha : s \trianglelefteq t\big\} \subset ([0,1]^\alpha)^2,$$

is a set of uniqueness for $\pi \in \Pi(m,m)$ for all $m \in \mathcal{P}([0,1]^\alpha)$.

Proof. The proof will be done by induction over α.

Step 0. First of all, if $\alpha = 1$, then the result follows from the observation that $B_1 = \{(s,t) : s \leq t\}$ is a set of uniqueness by elementary computations.

Step 1. Assume that $B_\alpha := \big\{(s,t) : s \trianglelefteq t\big\} \subset ([0,1]^\alpha)^2$ is a set of uniqueness, and consider the set $B_{\alpha+1}$. By the definition of lexicographic ordering, $(P_{\gamma\leq\alpha} \otimes P_{\gamma\leq\alpha})(B_{\alpha+1}) = B_\alpha$, so that we can write by the Disintegration Theorem

$$m_\alpha := (P_{\gamma\leq\alpha})_\sharp m, \quad m = \int m_t m_\alpha(dt), \;\pi = \int \big(\delta_{(t,t)} \otimes \pi_t\big) m_\alpha(dt),$$

with $m_t \in \mathcal{P}([0,1])$, $\pi_t \in \mathcal{P}([0,1]^2)$. We have used the fact that

$$\pi_\alpha := (P_{\gamma\leq\alpha} \otimes P_{\gamma\leq\alpha})_\sharp \pi \in \Pi(m_\alpha, m_\alpha),$$

and then the uniqueness property of B_α implies that $\pi_\alpha = (\mathbb{I},\mathbb{I})_\sharp m_\alpha = \int \delta_{(t,t)} m_\alpha(dt)$.

Note now that for m_α-a.e. $t \in [0,1]^\alpha$ it holds $\pi_t \in \Pi(m_t, m_t)$ and that

$$B_{\alpha+1} \cap \big(P^{-1}_{\gamma\leq\alpha}, P^{-1}_{\gamma\leq\alpha}\big)(t) = \Big\{(s,s') : P_{\gamma\leq\alpha}(s) = P_{\gamma\leq\alpha}(s') = t, P_{\alpha+1}(s) \leq P_{\alpha+1}(s')\Big\} = \{(t,t) \times B_1\}.$$

This is clearly a set of uniqueness, by Step 0, so that $\pi_t = (\mathbb{I},\mathbb{I})_\sharp m_t$. We thus conclude that also $B_{\alpha+1}$ is a set of uniqueness.

Step 2. Let $\alpha \in \omega_1$ be a limit ordinal. Then for all $\beta < \alpha$ the set $B_\beta = P_{\gamma\leq\beta}(B)$ is a set of uniqueness. Using the fact that

$$\big\{(s,t) \in [0,1]^\alpha \times [0,1]^\alpha : s = t\big\} = \bigcap_{\beta<\alpha} \big\{(t,s) : P_{\gamma\leq\beta}(t) = P_{\gamma\leq\beta}(s)\big\},$$

and observing that

$$(P_{\gamma\leq\beta} \otimes P_{\gamma\leq\beta})_\sharp \pi = (P_{\gamma\leq\beta}, P_{\gamma\leq\beta})_\sharp m,$$

we conclude that $\pi(\{s = t\}) = 1$, i.e. B is a set of uniqueness.

Step 3. By transfinite induction, we conclude that for every $\alpha \in \omega_1$ the set $B = \{(s,t) : s \trianglelefteq t\} \subset ([0,1]^\alpha)^2$ is a set of uniqueness. □

As noticed concerning the diagonal in the discussion before Assumption 2, $\{s \trianglelefteq t\}$ does not belong to the product σ-algebra if $\alpha = \omega_1$, so that the uniqueness question for $\alpha = \omega_1$ is meaningless.

Lemma 4.14. *Let $B \in \Theta_{m\otimes m}$ be the graph of a linear preorder $\preccurlyeq$ and $m \in \mathcal{P}([0,1])$ a probability measure such that for some $\kappa \in [0,1]$*

$$m(B^{-1}(x)) = \kappa \quad \textit{for } m\textit{-a.e. } x \in [0,1]. \tag{4.4}$$

Then m is concentrated on an equivalence class E_0 for $E = B \cap B^{-1}$ and $k = 1$. If (4.4) holds for all x

$$\forall x \in E_0, \forall y \in [0,1] \quad x \preccurlyeq y. \tag{4.5}$$

See Example C.12 for an example where the assumption $B \in \Theta_{m\otimes m}$ is not satisfied and the thesis is false.

Proof. Step 1. Let us prove the thesis for $\kappa = 1$. In this case, we do not need the assumption that the preorder B is linear.

By Fubini Theorem it follows that

$$m \otimes m(B^{-1}) = \int m(B^{-1}(x)) m(dx) = 1.$$

Thus $m \otimes m$ is concentrated on B^{-1}. Using the formula

$$B^{-1}(x) \cap (B \setminus E)(x) = \{y : y \preccurlyeq x\} \cap \{y : y \succ x\} = \emptyset$$

one has that $B^{-1} \cap (B \setminus E) = \emptyset$, and then $m \otimes m(B \setminus E) \leq m \otimes m([0,1] \setminus B^{-1}) = 0$.

Since $m \otimes m$ is invariant for the reflection w.r.t. the diagonal, $m \otimes m(B^{-1} \setminus E) = m \otimes m(B \setminus E) = 0$, finding that $m \otimes m$ is concentrated on E.

Again by Fubini Theorem,

$$m \otimes m = \int m_{\llcorner E(x)} m(dx) \quad \Longrightarrow \quad m(E(x)) \neq 0 \;\; m\text{-a.e. } x,$$

which implies that the image measure w.r.t. E is purely atomic.

Under the present assumption that $\kappa = 1$ there can be at most one equivalence class with positive measure: denoting this class by E_0, clearly $m(E_0) = 1$ and (4.5) holds if $m(B^{-1}(x)) = 1$ for all x.

Step 2. Notice first how the assumption that the preorder is linear implies $\kappa > 0$. Indeed,

$$B \cup B^{-1} = [0,1]^2$$

yields $m \otimes m(B) + m \otimes m(B^{-1}) \geq 1$. Since $m \otimes m$ is invariant for the reflection w.r.t. the diagonal, $m \otimes m(B) = m \otimes m(B^{-1})$, and then $2 \cdot m \otimes m(B) \geq 1 > 0$. By Fubini and (4.4) we conclude

$$\kappa = m \otimes m(B) > 0.$$

Let $X = B^{-1}(\bar{x})$ such that $m(B^{-1}(\bar{x})) = \kappa$. For m-a.e. $x \succ \bar{x}$ $B^{-1}(\bar{x}) \subset B^{-1}(x)$ and by (4.4) one has

$$m_{\llcorner [0,1]\setminus X}(B^{-1}(x)) = m(B^{-1}(x) \setminus B^{-1}(\bar{x})) = 0.$$

If $m([0,1] \setminus X) = 1 - \kappa > 0$ we would reach an absurd, as (4.4) would hold with $\kappa = 0$ for the probability measure $\tilde{m} := (m_{\llcorner [0,1]\setminus X})/(1-\kappa)$. This yields $\kappa = 1$, proving the thesis by the first step. □

We can now prove the main theorem of the section.

Proof of Theorem 4.9. The following steps of the proof show that given any measure $\mu \in \mathcal{P}([0,1])$ there exists an ordinal $\bar\alpha \in \omega_1$ and an order preserving Borel map

$$h_{\bar\alpha} \quad : \quad ([0,1], \preccurlyeq) \quad \rightarrow \quad ([0,1]^{\bar\alpha}, \trianglelefteq)$$

such that

$$\mu = \int \mu_t m(dt) \qquad m := (h_{\bar\alpha})_\sharp \mu$$

with $\mu_{h_{\bar\alpha}(x)}$ concentrated on the single equivalence class $E(x) \subset h_{\bar\alpha}^{-1}(h_{\bar\alpha}(x))$, defined in (4.2), for μ-a.e. x. We prove that if one removes a μ-negligible set $h_{\bar\alpha}$ is the quotient projection w.r.t. E.

Since the map is order preserving, the image set B' of B in the quotient space is clearly a subset of $\big\{(s,t) \in [0,1]^\alpha \times [0,1]^\alpha : s \trianglelefteq t\big\}$. The fact that B' is a set of uniqueness of $\Pi(m,m)$, for all m Borel, follows then immediately from Lemma 4.13, proving the last part of the theorem.

Step 1. Define the function

$$[0,1] \quad \ni \quad x \quad \mapsto \quad h_1(x) := \mu(\{z : z \preccurlyeq x\}) = \mu(B^{-1}(x)) \quad \in \quad [0,1].$$

Since B is Borel, by Lemma 4.10 h_1 is Borel and

$$x \preccurlyeq y \quad \Longrightarrow \quad h_1(x) \leq h_1(y).$$

Step 2. Let $\alpha \in \omega_1$ and assume that there exists a μ-measurable order preserving map $h_\alpha : [0,1] \to [0,1]^\alpha$, where $[0,1]^\alpha$ is ordered by the lexicographic ordering $\trianglelefteq$.

Since $[0,1]^\alpha$ is Polish because $|\alpha| \leq \aleph_0$, then the disintegration

$$\mu = \int \mu_t (h_\alpha)_\sharp (dt)$$

is well defined and strongly consistent.

By redefining μ_t on a set of measure 0 w.r.t. $(h_\alpha)_\sharp \mu$, we can assume that $[0,1]^\alpha \ni t \mapsto \mu_t \in \mathcal{P}([0,1])$ is Borel, so that the map

$$[0,1] \times [0,1]^\alpha \quad \ni \quad (x,t) \quad \mapsto \quad \mu_t(B^{-1}(x)) \quad \in \quad [0,1]$$

is Borel by Lemma 4.12 and the Borel Isomorphism Theorem (Theorem 3.3.13, page 99 of [19]).

Step 3. Consider the function

$$\begin{array}{rccc} h_{\alpha+1} & : & [0,1] & \rightarrow & [0,1]^{\alpha+1} \\ & & x & \mapsto & h_{\alpha+1}(x) := (h_\alpha(x), \mu_{h_\alpha(x)}(B^{-1}(x))) \end{array}$$

Since $t \mapsto h_\alpha(t)$ is Borel, then $(t,x) \mapsto \mu_{h_\alpha(t)}(B^{-1}(x))$ is Borel by Lemma 4.12, being the composition of two Borel maps. It is clearly order preserving if h_α is and $[0,1]^{\alpha+1}$ is ordered lexicographically. Note that $P_{\beta \leq \alpha} \circ h_{\alpha+1} = h_\alpha$.

Step 4. Assume that α is a limit ordinal, and that the Borel functions $h_\beta : [0,1] \to [0,1]^\beta$, $\beta < \alpha$, have been constructed in such a way that $P_{\gamma \leq \beta} \circ h_\delta = h_\beta$ for all $\beta \leq \delta$. Then we construct the Borel order preserving map

$$\begin{array}{rccc} h_\alpha & : & [0,1] & \rightarrow & [0,1]^\alpha \\ & & x & \mapsto & (P_\beta \circ h_\alpha)(x) := P_\beta(h_\beta(x)) \ \forall \beta < \alpha \end{array}$$

By transfinite induction, we can construct a Borel order preserving map $h : [0,1] \to [0,1]^{\omega_1}$, where $[0,1]^{\omega_1}$ is equipped with the Borel σ-algebra.

Step 5. Consider the family of Borel equivalence relations

$$\mathcal{E} := \Big\{ E_\alpha = \{x \sim y \ \Leftrightarrow \ h_\alpha(x) = h_\alpha(y)\}, \alpha \in \omega_1 \Big\}.$$

We observe that if $\{\alpha_n\} \subset \omega_1$ and $\alpha = \sup_n \alpha_n \in \omega_1$, then $\cap_n E_{\alpha_n} = E_\alpha$ by the definition of h_α. By Theorem A.11, there exists $\bar\alpha \in \omega_1$ such that the disintegration

$$\mu = \int \mu_t ((h_{\bar\alpha})_\sharp \mu)(dt)$$

is the sharpest one, in the sense that any other disintegration w.r.t. h_β, $\beta \in \omega_1$, namely

$$\mu = \int \mu_s ((h_\beta)_\sharp \mu)(ds)$$

can be written by means of

$$\mu\llcorner_{h_{\bar\alpha}^{-1}(\mathcal{B}([0,1]^{\bar\alpha}))} = \int r_s((h_\beta)_\sharp\mu)(ds), \quad \mu = \int \mu_s((h_\beta)_\sharp\mu)(ds) = \int \bigg(\int \mu_t r_s(dt) \bigg) ((h_\beta)_\sharp\mu)(ds).$$

We write here $\mu\llcorner_{\mathcal{C}}$ for the restriction of a measure to the σ-algebra $\mathcal{C} \subset \mathcal{B}$.

By (A.5), we conclude that there exists a Borel function $s : [0,1]^{\bar\alpha} \to [0,1]^\beta$ such that $(h_\beta)_\sharp\mu$ is concentrated on the graph of $s = s(t)$, where $t \in [0,1]^{\bar\alpha}$.

Step 6. From the definition of $h_{\bar\alpha+1}$ and Step 5, it follows that

$$\mu_t(B^{-1}(x)) = s(t) \quad \text{for } (h_{\bar\alpha})_\sharp\mu\text{-a.e. } t,\ \forall x \in h_{\bar\alpha}^{-1}(t),$$

i.e. $\mu_t(B^{-1}(x))$ is constant on $h_{\bar\alpha}^{-1}(t)$, for $(h_{\bar\alpha})_\sharp\mu$-a.e $t \in [0,1]^{\bar\alpha}$: in particular removing a μ-negligible saturated set we can assume that by Lemma 4.14 the measure μ_t is concentrated on the first equivalence class contained in $B(h_{\bar\alpha}^{-1}(t))$, denote it with $E_{h_{\bar\alpha}}$.

Step 7. Define the function $\bar h : [0,1] \to [0,1]^{\bar\alpha+1}$ as

$$\bar h(x) := \big(h_{\bar\alpha}(x), 1 - \mu_{h_{\bar\alpha}(x)}(B(x)) \big).$$

The $(\bar\alpha+1)$-component is 0 only on the unique class $E_{h_{\bar\alpha}(x)}$ of E on which $\mu_{h_{\bar\alpha}(x)}$ is concentrated. From the disintegration formula w.r.t. $h_{\bar\alpha}$ then

$$\mu\big(\bar h^{-1}(P_{\bar\alpha+1}^{-1}((0,1])) \big) = 0.$$

We conclude that, up to the above μ-negligible set, $h_{\bar\alpha}$ is the quotient map w.r.t. E and using the fact that $[0,1]^{\bar\alpha}$ is Polish

$$\mu = \int \mu_t((h_{\bar\alpha})_\sharp\mu)(dt)$$

is a strongly consistent disintegration for E. □

Corollary 4.15. *If B is a Borel linear order, then B is a set of uniqueness in $\Pi(m,m)$ for every $m \in \mathcal{P}([0,1])$.*

Proof. It is sufficient to observe that

$$E := B \cap B^{-1} = \{x = y\}.$$

Hence the map $h_{\bar\alpha} : [0,1] \to [0,1]^{\bar\alpha}$ constructed in the previous proof is order preserving and moreover

$$m = \int \delta_{x(t)}((h_{\bar\alpha})_\sharp m)(dt).$$

Removing the cross negligible set N where $h_{\bar\alpha} \otimes h_{\bar\alpha}$ is not invertible, we have that the uniqueness problem can be stated as a uniqueness problem in $[0,1]^{\bar\alpha}$ with the lexicographic ordering $\trianglelefteq$. By Lemma 4.13 uniqueness of π follows. □

Remark 4.16. In Theorem 4.9 and Corollary 4.15 one can assume that $B \in \Theta(\mu,\mu)$. The proof is analogous to the one above, but relies on the following lemma instead of Lemmas 4.10-4.12.

Lemma 4.17. *Let $B \in \Theta(m,m)$, with $m \in \mathcal{P}([0,1])$, and $h : ([0,1],\Theta_m) \to ([0,1],\mathcal{B})$ a measurable map. Let $m = \int m_t \xi(dt)$ be the disintegration of m w.r.t. h. Then*

1. *the map $x \mapsto m_t(B(x))$ is m-measurable for ξ-a.e. t;*
2. *the map $(x,t) \mapsto m_t(B(x))$ is π-measurable for $\pi \in \Pi(m,\xi)$;*
3. *the map $x \mapsto m_{h(x)}(B(x))$ is m-measurable.*

For simplifying the notation, in the following we will denote the disintegration of $\eta \in \mathcal{P}(X_1 \times X_2)$ w.r.t. the projection P_1 as

$$\int \phi\eta = \int \bigg(\int \phi(x_1,x_2)\eta_{x_1}(dx_2) \bigg) ((P_1)_\sharp\eta)(dx_1).$$

Proof. Step 1: Point (1). By the Disintegration Theorem A.7 $(x,t) \mapsto m_t(B^{-1}(x))$ is $m \otimes \xi$-measurable, where $\xi := (h_\alpha)_\sharp m$:

$$\begin{aligned} m \otimes m(B^{-1}) &= \int m(B^{-1}(x)) m(dx) \\ &= \int \bigg\{ \int m_t(B^{-1}(x)) \xi(dt) \bigg\} m(dt) = \int m_t(B^{-1}(x)) (m \otimes \xi)(dx, dt). \end{aligned}$$

By Fubini Theorem $x \to m_t(B(x))$ is therefore m-measurable for $t \in [0,1] \setminus N$, $\xi(N) = 0$.

Step 2: Point (2). Let $\hat{\pi} \in \Pi(m, \xi)$ and $\hat{\pi} = \int \hat{\pi}_t \xi(dt)$ its disintegration w.r.t. the projection on the second variable. Define the Borel measure on $[0,1]^3$

$$\eta := \int m_t \hat{\pi}(dxdt) = \int m_t \otimes \hat{\pi}_t \xi(dt).$$

Notice that $(P_1)_\sharp \eta = \int m_t \xi(dt) = m$. Similarly, $(P_2)_\sharp \eta = (P_1)_\sharp \hat{\pi} = m$. In particular, $(P_{12})_\sharp \eta \in \Pi(m,m)$. Since $(P_{12})^{-1}(\Theta(m,m)) \subset \Theta_\eta$, then $B \times [0,1] \in \Theta_\eta$. By the Disintegration Theorem one finds that $(x,t) \mapsto m_t(B(x))$ is $\hat{\pi}$-measurable and

$$\eta(B \times [0,1]) = \int m_t(B(x)) \hat{\pi}(dxdt).$$

Step 3: Point (3). Finally by taking $\pi := (\mathbb{I}, h)_\sharp m \in \Pi(m, \xi)$ one has the last point of the statement. □

Hence, our sufficient condition for uniqueness is the following:

Sufficient condition for uniqueness: A' is a subset of a linear order $B \in \Theta(m,m)$.

As the diagonal is A'-acyclic, and therefore by the Axiom of Choice it can be completed to a linear order, this is again a measurability assumption.

An easy case is covered by the next lemma.

Lemma 4.18. *If A' is acyclic and m purely atomic, then A' is a subset of a Borel linear order on $[0,1]$ and hence a set of uniqueness.*

Proof. If m is purely atomic with atoms on $\{\alpha_i\}_{i\in\mathbb{N}}$, it is enough to prove that we can find a linear order on the set $\{(\alpha_i, \alpha_j), i, j \in \mathbb{N}\}$. In fact it is fairly easy to extend its graph R to a Borel linear order on $[0,1]$ by defining

$$\alpha \preccurlyeq \beta \quad \Longleftrightarrow \quad \begin{cases} \alpha R \beta & \alpha, \beta \in \{\alpha_i\}_{i\in\mathbb{N}} \\ \alpha \in \{\alpha_i\}_{i\in\mathbb{N}} & \beta \notin \{\alpha_i\}_{i\in\mathbb{N}} \\ \alpha \leq \beta & \alpha, \beta \notin \{\alpha_i\}_{i\in\mathbb{N}} \end{cases}$$

Now, every acyclic set A' on $\{(\alpha_i, \alpha_j)\}_{i,j\in\mathbb{N}}$ containing $\{(\alpha_i,\alpha_i), i \in \mathbb{N}\}$ defines a partial order relation by setting

$$\alpha_i \preccurlyeq \alpha_j \quad \Longleftrightarrow \quad (\alpha_i, \alpha_j) \in A'.$$

By countably many operations one can complete this partial order into a linear one. □

An example of a set A for which A' is a set of uniqueness is presented in Figure 2. By setting

$$c(x,y) = \begin{cases} 1 & \Gamma \\ 0 & A \setminus \Gamma \\ +\infty & [0,1]^2 \setminus A \end{cases}$$

the uniqueness of the transport plan in A is related to a problem of optimality.

Remark 4.19. We observe here that given an A-acyclic set Γ, and assuming for simplicity that $P_1(\Gamma) = [0,1]$, one can define a preorder on $[0,1]$ by

$$x \preccurlyeq x' \quad \Longleftrightarrow \quad \exists \{(x_i, y_i)\}_{i=0,\dots,I} \subset \Gamma,\ (x_{i+1}, y_i) \in A,\ x_0 = x, x_{I+1} = x',$$

i.e. x, x' are connected by an axial path. The equivalence relation E defined in (4.2) corresponds to the axial equivalence relation (3.4), so that we can state equivalently that if $\preccurlyeq$ can be extended to a linear preorder, then A is a set of uniqueness.

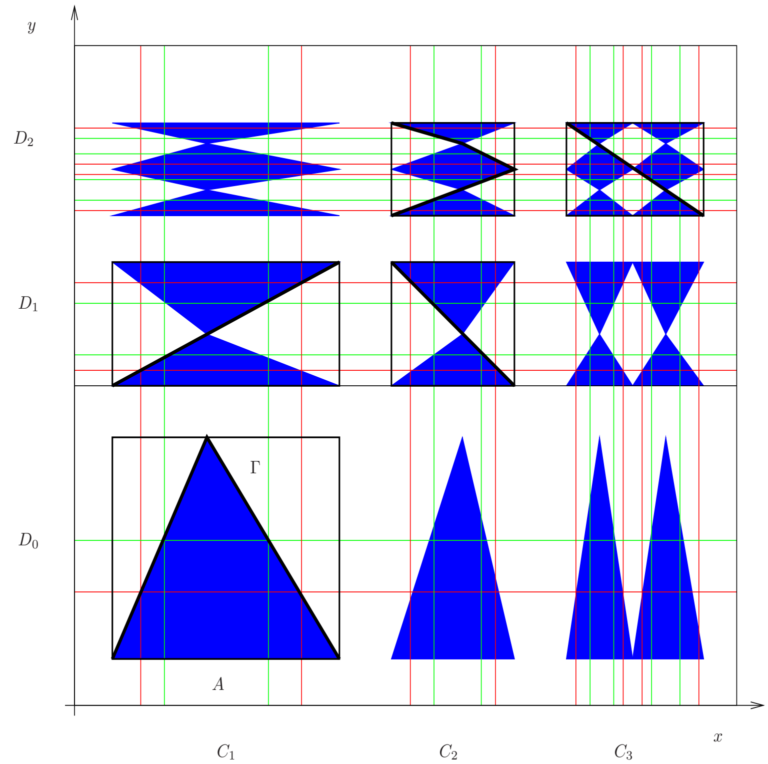


Figure 2: The set where A should be contained in order to have that A' is a the graph of the standard order $\leq$ on $[0,1]$. The bold curves are the limbs of Γ, and two axial path are represented.

5. Optimality

The last problem we want to address is the problem of optimality of a measure $\pi \in \Pi(\mu,\nu)$ w.r.t. the functional $\mathcal{I}$ defined in (1.2). We recall that a plan $\pi \in \Pi(\mu,\nu)$ is said to be optimal if

$$\mathcal{I}(\pi) = \int c(x,y)\pi(dxdy) = \min_{\tilde{\pi}\in\Pi(\mu,\nu)} \mathcal{I}(\tilde{\pi}).$$

In this section the function c is assumed to be a positive Π^1_1-function.

Definition 5.1 (Cyclical monotonicity)**.** A subset Γ of $[0,1]^2$ is *c-cyclically monotone* when for all I, $i = 1,\dots,I$, $(x_i,y_i) \in \Gamma$, $x_{I+1} := x_1$ we have

$$\sum_{i=1}^{I} \big[c(x_{i+1},y_i) - c(x_i,y_i)\big] \geq 0.$$

A transference plan $\pi \in \Pi(\mu,\nu)$ is *c-cyclically monotone* if there exists a c-cyclically monotone set Γ such that $\pi(\Gamma) = 1$.

As usual, by inner regularity and by the fact that for π fixed c coincides with a Borel function up to a negligible set, the set Γ for that given measure π can be taken σ-compact and $c\llcorner_\Gamma$ Borel.

We recall that a necessary condition for being optimal is that the measure is concentrated on a c-cyclically monotone set. A proof is provided for completeness in Proposition B.16.

Lemma 5.2. *If π is optimal, then it is c-cyclically monotone.*

Having a necessary condition which gives some structure to the problem, we have to specify a sufficient condition which should be tested in each equivalence class. We list some important remarks.

(1) The optimality is implied by the fact that there exists a sequence of functions $\phi_n \in \mathcal{L}^1(\mu)$, $\psi_n \in \mathcal{L}^1(\nu)$ such that $\phi_n(x) + \psi_n(y) \leq c(x,y)$ and

$$\int \phi_n \mu + \int \psi_n \nu = \int (\phi_n + \psi_n)\pi \nearrow \int c\pi.$$

(2) For l.s.c. costs or costs satisfying $c(x,y) \leq f(x) + g(y)$, $f \in L^1(\mu)$ and $g \in L^1(\nu)$-measurable, the converse of Point (1) holds.

(3) Another condition is that there is an optimal pair $\phi, \psi : [0,1] \to [-\infty, +\infty)$, respectively μ-measurable and ν-measurable, such that $\phi(x) + \psi(y) \leq c(x,y)$ for all $(x,y) \in [0,1]^2$ and $\phi(x) + \psi(y) = c(x,y)$ π-a.e..

For completeness we prove the sufficiency of the last condition, proved also in [4].

Lemma 5.3. *Suppose there exist Borel functions $\phi, \psi : [0,1] \to [-\infty, +\infty)$ and $\Gamma \subset [0,1]^2$ such that*

$$\begin{aligned} &\phi(x) + \psi(y) < c(x,y) && \forall (x,y) \in [0,1]^2 \setminus \Gamma \\ &\phi(x) + \psi(y) = c(x,y) && \forall (x,y) \in \Gamma. \end{aligned}$$

If $\exists \pi \in \Pi^f(\mu,\nu)$ such that $\pi(\Gamma) = 1$, then

$$\pi \in \Pi(\mu,\nu) \ \textit{optimal} \quad \Longleftrightarrow \quad \pi(\Gamma) = 1.$$

It is trivial to extend the proposition to the case of $\phi : [-\infty, +\infty)$ μ-measurable and $\psi : [-\infty, +\infty)$ ν-measurable, just redefining the functions on negligible sets in order to be Borel.

Proof. Let $\bar{\pi}$ be an optimal transference plan and $\pi \in \Pi^f(\mu,\nu)$ concentrated on Γ. Hence μ and ν are concentrated on the sets $\{\phi > -\infty\}$, $\{\psi > -\infty\}$ respectively.

Step 1. We prove that if $\lambda \in \Lambda$ and $\psi\lambda$, $(\phi + \psi)\lambda$ are Borel measures (assuming eventually the value ∞) concentrated on $\{\phi + \psi > -\infty\}$, then $\int \{\phi + \psi\}\lambda = 0$.
Since $(\phi + \psi)\lambda$ is a Borel measure, one can consider the following integrals

$$\int_{[0,1]^2} \{\phi + \psi\}\lambda = \lim_{M \to \infty} \int_{\{|\phi| < M\}} \{\phi + \psi\}\lambda.$$

Since also $\psi\lambda$ is a Borel measure and $\phi\chi_{|\phi|<M}\lambda \in \mathcal{M}([0,1]^2)$,

$$\begin{aligned} \lim_{M \to \infty} \int_{\{|\phi| < M\}} \{\phi + \psi\}\lambda &= \lim_{M \to \infty} \left\{ \int_{\{|\phi| < M\}} \phi\lambda + \int_{\{|\phi| < M\}} \psi\lambda \right\} \\ &\overset{\lambda \in \Lambda}{=} \lim_{M \to \infty} \int_{\{|\phi| < M\}} \psi\lambda = \int \psi\lambda = \lim_{M \to \infty} \int_{\{|\psi| < M\}} \psi\lambda = 0. \end{aligned}$$

Step 2. Let $\lambda := \bar{\pi} - \pi$, with $\bar{\pi} \in \Pi^f(\mu,\nu)$.
Define $\phi_M := (\phi \wedge M) \vee (-M)$ and $\psi_M := (\psi \wedge M) \vee (-M)$: it is immediate to verify that from $\phi + \psi = c$ on Γ

$$\Gamma(\{\phi \leq M\}) \subset \{\psi \geq -M\}, \quad \Gamma^{-1}(\{\psi \leq M\}) \subset \{\phi \geq -M\},$$

and then

$$\phi_M(x) + \psi_M(y) \leq c(x,y), \quad \phi_M(x) + \psi_M(y) \geq 0 \text{ on } \Gamma.$$

In particular, $\phi_M\lambda$ and $(\phi_M + \psi_M)\lambda$ are finite Borel measures.

Since ϕ_M, ψ_M converge pointwise, then $\phi_M + \psi_M$ converges to c in $L^1(\pi)$, yielding immediately

$$\int_{[0,1]^2} c\lambda \geq \lim_M \int_{[0,1]^2} \{\phi_M + \psi_{-M}\}\lambda.$$

The r.h.s. vanishes by Step 1, showing the optimality of π:

$$0 \geq \mathcal{I}(\bar{\pi}) - \mathcal{I}(\pi) = \int_{[0,1]^2} c\lambda \geq \lim_M \int_{[0,1]^2} \{\phi_M + \psi_{-M}\}\lambda = 0.$$

□

From the formulas

$$\phi(x,\bar x,\bar y) = \inf\left\{ \sum_{i=0}^{I} c(x_{i+1},y_i) - c(x_i,y_i), (x_i,y_i) \in \Gamma \text{ finite}, (x_0,y_0) = (\bar x,\bar y),\ x_{I+1} = x \right\}, \tag{5.2a}$$

$$\psi(y,\bar x,\bar y) = c(x,y) - \phi(x,\bar x,\bar y), \quad (x,y) \in \Gamma, \tag{5.2b}$$

it is always possible to construct an optimal couple ϕ, ψ in an analytic subset of Γ containing $(\bar x,\bar y)$ such that $(-\phi,\psi)$ are Σ^1_1-functions. In Remarks C.3, C.6 it is shown that ϕ, ψ are $\mathcal{A}$-functions, but using the facts that

$$g \in \Pi^1_1(X \times Y) \ \Rightarrow\ \inf_y g(x,y) \in \Pi^1_1(X), \quad c\llcorner_\Gamma \text{ is Borel},$$

the above better estimate follows.

In Remark 5.10 we show that ϕ defines a natural linear preorder on $[0,1]$, and that we can state a particularly concise condition. In Section C, instead, the idea of extending $\phi(x,\bar x,\bar y)$, $\psi(y,\bar x,\bar y)$ to larger sets is developed in a general framework.

Here we consider the easiest equivalence relation for which the procedure at Page 8 can be applied. This equivalence relation has been also used in [3].

Definition 5.4 (Closed cycles equivalence relation)**.** We say that $(x,y)\bar E(x',y')$ or (x,y) *is equivalent to* (x',y') *by closed cycles* if there is a closed cycle with finite cost passing through them: there are $(x_i,y_i) \in \Gamma$ such that $(x_0,y_0) = (x,y)$ and $(x_j,y_j) = (x',y')$ for some $j \in \{0,\dots,I\}$ such that

$$\sum_{i=1}^{I} c(x_i,y_i) + c(x_{i+1},y_i) < +\infty, \qquad x_{I+1} := x_0.$$

It is easy to show that this is an equivalence relation, and it follows directly from (5.2) or the analysis of Section C that in each equivalence class there are optimal potentials ϕ, ψ.

Lemma 5.5. *The equivalence relation $\bar E$ satisfies the following.*

1. *Its equivalence classes are in Σ^1_1.*
2. *It satisfies the crosswise structure* (2.2)*.*

The above lemma can be seen as a straightforward consequence of Lemma C.5 and Corollary C.7 of Section C. Since it is elementary, we give here a direct proof.

Proof. For the Point (1), just observe that for all $I \in \mathbb{N}$

$$\sum_{i=0}^{I} c(x_i,y_i) + c(x_{i+1},y_i), \quad x_{I+1} = x_0$$

is a Π^1_1-function, so that

$$Z_I(\bar x,\bar y) = \left\{ (x_1,y_1,\dots,x_I,y_I) \in \Gamma^I : \sum_{i=1}^{I} c(x_i,y_i) + c(x_{i+1},y_i) + c(\bar x,\bar y) + c(x_1,\bar y) < +\infty, x_{I+1} = \bar x \right\}$$

is in Σ^1_1.

The equivalence class of $(\bar x,\bar y)$ is then given by

$$\bigcup_{I\in\mathbb{N}} \bigcup_{i=1}^{I} P_{2i-1,2i}(Z_I) \in \Sigma^1_1,$$

where we used the fact that Σ^1_1 is closed under projection and countable union (see Appendix B.1 or Chapter 4 of [19]).

The proof of Point (2) follows from the straightforward observation that $(x,y)\bar E(x',y)$ and $(x,y)\bar E(x,y')$ whenever $(x,y),(x',y),(x,y') \in \Gamma$: just consider the closed cycle with finite cost made of the two points $(x_0,y_0) := (x,y)$ and $(x_1,y_1) := (x',y)$, or $(x_1,y_1) := (x,y')$. □

Let now $\pi \in \Pi^f(\mu,\nu)$ be a c-cyclically monotone transference plan, and let Γ be a c-cyclically monotone set where π is concentrated. Let $\bar E$ be the equivalence class of Definition 5.4.

As in the previous section, by Lemmas 5.3 and 2.5 non optimality can occur because of two reasons:

(1) either the disintegration is not strongly consistent,
(2) or the push forward of some measure $\pi' \in \Pi^{\rm opt}(\mu,\nu)$ is not supported on the diagonal in the quotient space.

In the next section we give examples which show what can happen when one of the two situations above occurs. Here we conclude with two results, which yield immediately the optimality of π.

Let h_X, h_Y be the quotient maps. By redefining them on a set of measure 0, the condition of strong consistency implies that h_X, h_Y can be considered as Borel maps with values in $[0,1]$. In particular, the set

$$A' := (h_X \otimes h_Y)(\{c < +\infty\}) \tag{5.3}$$

is analytic. Note that

$$(h_X \otimes h_Y)_\sharp \tilde\pi(A') = 1 \quad \forall \tilde\pi \in \Pi^f(\mu,\nu),$$

i.e. the transport plans with finite cost are concentrated on A', and moreover for the π under consideration

$$(h_X \otimes h_Y)_\sharp \pi = (\mathbb{I},\mathbb{I})_\sharp m,$$

where $m = (h_X)_\sharp \mu = (h_Y)_\sharp \nu$ by Lemma 5.5 and Lemma 2.4.

Theorem 5.6. *Assume that the disintegration w.r.t. the equivalence relation $\bar E$ is strongly consistent. If A' is a set of uniqueness in $\Pi(m,m)$, then π is optimal.*

The proof is a simple application of Lemma 2.5.

The next corollary is a direct consequence of Lemma 4.18.

Corollary 5.7. *If $m = (h_X)_\sharp \mu$ is purely atomic, then the c-cyclical monotone measure π is optimal.*

We now give a simple condition which implies that the image measure m is purely atomic.

Proposition 5.8. *Assume that π satisfies the following assumption: there exists a countable family of Borel sets $A_i, B_i \subset X$, $i \in \mathbb{N}$ such that*

$$\pi\bigg(\bigcup_i A_i \times B_i\bigg) = 1, \tag{5.4a}$$

and

$$\mu \otimes \nu\big(\cup_i (A_i \times B_i) \cap \{c = +\infty\}\big) = 0. \tag{5.4b}$$

Then the image measure m is purely atomic.

Proof. First of all, we can assume that $\Gamma \subset \cup_i A_i \times B_i$, where Γ is a c-cyclically monotone set such that $\pi(\Gamma) = 1$.

Step 1. The assumption (5.4b) and Fubini Theorem imply that there is $\bar x_i \in A_i$ such that

$$\bar B_i := P_2\big((A_i \times B_i \cap \{c < +\infty\})_{\bar x_i}\big),$$

has full $\nu \llcorner_{B_i}$-measure, where for $C \subset [0,1]^2$ we define

$$C_x := C \cap \{x\} \times [0,1].$$

Then there is $\bar y_i \in \bar B_i$ such that

$$\bar A_i := P_2\big((A_i \times B_i \cap \{c < +\infty\})_{\bar y_i}\big)$$

has full $\mu \llcorner_{A_i}$-measure. The functions ϕ, ψ given by formula (5.2) starting from $(\bar x_i, \bar y_i)$ provide then optimal potentials on the sets $\bar A_i \times \bar B_i$.

Step 2. It is now fairly easy to show that the new sets $\bar A_i$, $\bar B_i$, $i \in \mathbb{N}$, satisfy again conditions (5.4), and that in $\Gamma \cap \cup_i \bar A_i \times \bar B_i$ each equivalence class for the closed cycles equivalence relation contains at least one $\bar A_i \times \bar B_i$. Then it follows that m is purely atomic. □

The case of Point (3) of Page 3 corresponds to a single global class.

Remark 5.9. Let Γ be a c-cyclically monotone set where π is concentrated. The proof shows actually that in each set $\Gamma \cap (A_i \times B_i)$

$$\phi(x) + \psi(y) = c(x,y)$$

up to a cross-negligible set. This is clearly a stronger condition than $c\llcorner_\Gamma < +\infty$ $\mu \otimes \nu$-a.e..

Remark 5.10. From the definition of the optimal couple $(\phi(\cdot,\bar{x},\bar{y}),\psi(\cdot,\bar{x},\bar{y}))$, we can define the following relation on $P_1(\Gamma)$.

Definition 5.11. We say that $x \geq_c x'$ if $\exists y \in \Gamma(x)$ such that $\phi(x',x,y) < +\infty$: equivalently there are points $(x_i,y_i) \in \Gamma$, $i = 0,\dots,I$ such that $x_0 = x$, $x_{I+1} = x'$ and $\sum_i c(x_{i+1},x_i) + c(x_i,y_i) < +\infty$.

When we consider as uniqueness condition the Borel linear order condition of page 18, the results of this section can be rephrased as the fact that $\leq_c$ can be completed into a Borel linear preorder $\leq_c$ such that

(1) $x \leq_c x$;
(2) for all $x,x' \in [0,1]$, at least $x \leq_c x'$ or $x' \leq_c x$.
(3) $x \leq_c x'$ and $x' \leq_c x$ implies that they belong to a closed cycle with finite cost.

In fact, by using Theorem 4.9 it follows that the disintegration w.r.t. the equivalence classes is strongly supported and that the image set A' is a set of uniqueness: just observe that if $x \leq_c x'$, then there is an axial path connecting them, so that A' is contained in the graph of a Borel linear order on $[0,1]$.

5.1. **Extension of the construction.** The approach we are proposing can be generalized as follows.

Assumption 5. Assume that for any $(\bar{x},\bar{y}) \in \Gamma$ there exist universally measurable subsets $A_{(\bar{x},\bar{y})}$, $B_{(\bar{x},\bar{y})}$ of $[0,1]$ and universally measurable functions $\phi_{(\bar{x},\bar{y})}$, $\psi_{(\bar{x},\bar{y})}$ satisfying

$$(\bar{x},\bar{y}) \in A_{(\bar{x},\bar{y})} \times B_{(\bar{x},\bar{y})} \tag{5.5a}$$

$$\phi_{(\bar{x},\bar{y})}(x) + \psi_{(\bar{x},\bar{y})}(y) \leq c(x,y) \qquad \forall (x,y) \in A_{(\bar{x},\bar{y})} \times B_{(\bar{x},\bar{y})} \tag{5.5b}$$

$$\phi_{(\bar{x},\bar{y})}(x) + \psi_{(\bar{x},\bar{y})}(y) = c(x,y) \qquad \forall (x,y) \in A_{(\bar{x},\bar{y})} \times B_{(\bar{x},\bar{y})} \cap \Gamma. \tag{5.5c}$$

We can define the relation R

$$(x,y)R(x',y') \quad \Longleftrightarrow \quad (x',y') \in A_{(x,y)} \times B_{(x,y)}.$$

Assume that there exist partitions $\{X_\alpha\}_\alpha$, $\{Y_\alpha\}_\alpha$ of $[0,1]$ such that each $X_\alpha \times Y_\alpha \subset A_{(x_\alpha,y_\alpha)} \times B_{(x_\alpha,y_\alpha)}$ for some $(x_\alpha,y_\alpha) \in \Gamma$. Then optimality holds if the equivalence relation induced by $\{X_\alpha \times Y_\beta\}_{\alpha,\beta}$ satisfies Assumptions 1, 2, 3, i.e. if the disintegrations w.r.t. $\{X_\alpha \times Y_\beta\}_{\alpha,\beta}$ is strongly consistent, $\pi(\cup_\alpha X_\alpha \times Y_\alpha) = 1$ and A' of (5.3) is a set of uniqueness.

A method for constructing a relation R satisfying Assumption 5 and the crosswise condition w.r.t. Γ (Definition 2.3) is exploited in Appendix C.

6. Examples

In this section we study the dependence of our construction w.r.t. the choice of Γ, and the necessity of the assumptions in Theorem 5.6.

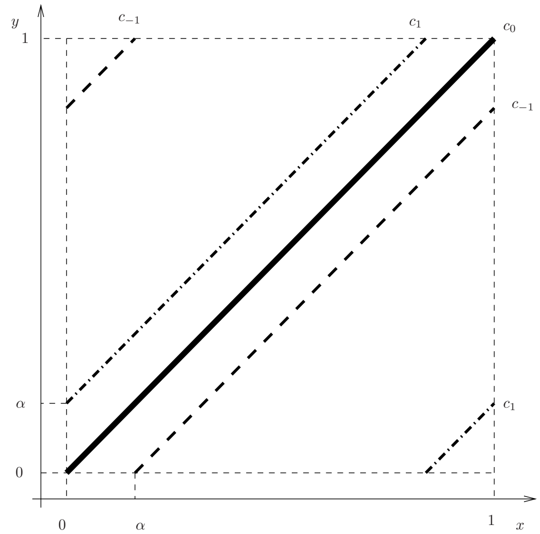
6.1. **Dependence w.r.t. the set Γ.** We consider the situation where the assumptions of Theorem 5.6 do not hold, so that either we do not have the strong consistency of the disintegration, or the set A' is not a set of uniqueness. Keeping fixed μ, ν, c and the plan $\pi \in \Pi(\mu,\nu)$, varying Γ, the following cases are possible:

(1) Strong consistency of the disintegration is not satisfied for any choice of Γ, and the plan we are testing can either be optimal or not (Example 6.1, Example 6.2).
(2) Strong consistency can be satisfied or not, depending on Γ, and, when it is, the quotient problem can be both well posed (A' is a set of uniqueness) or not (Example 6.3, Example 6.2). We are testing an optimal plan.
(3) Strong consistency is always satisfied, but the image measure m is not atomic (Example 6.4). The plan we are testing can either be optimal or not.

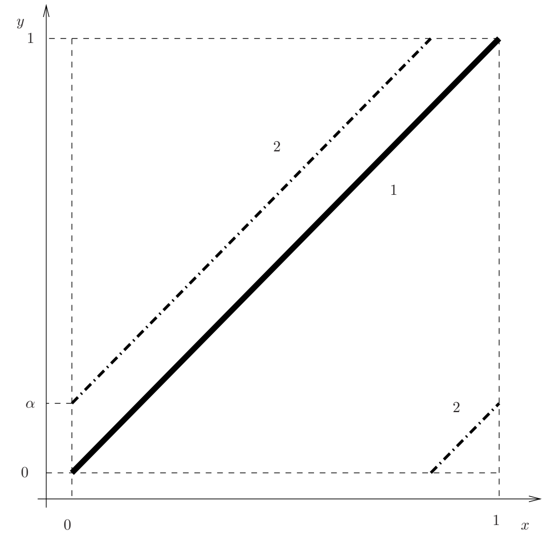
In Figure 3, for each example we draw the pictures of the set in $[0,1]^2$ where c is finite.

Example 6.1. Consider $\mu = \nu = \mathcal{L}^1$ with the cost given by

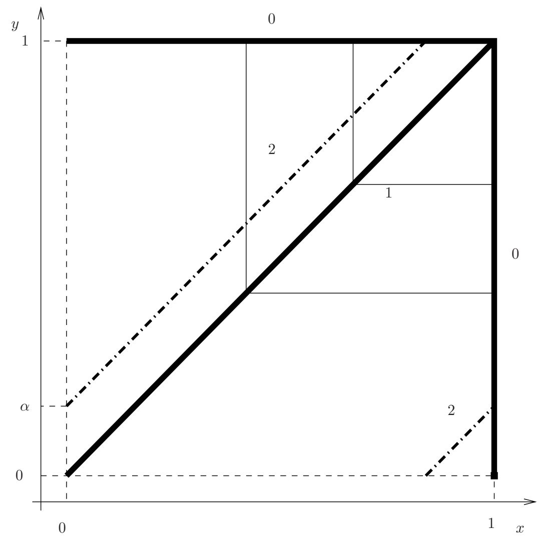
$$c(x,y) = \begin{cases} c_0 & y - x = 0 \\ c_1 & y - x = \alpha \pmod 1 \\ c_{-1} & y - x = -\alpha \pmod 1 \\ +\infty & \text{otherwise} \end{cases} \qquad \text{with } \alpha \in [0,1] \setminus \mathbb{Q} \text{ and } c_1 + c_{-1} \geq 2c_0.$$



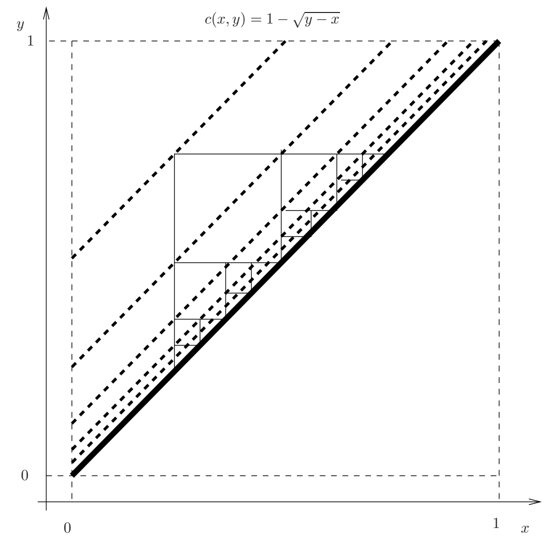
(a) When $\alpha \notin \mathbb{Q}$, the cycle decomposition of the plan $\pi = (\mathrm{Id}, \mathrm{Id})_\sharp \mathcal{L}^1$ gives always a non-measurable disintegration (Ex. 6.1).



(b) Disintegration sometimes only consistent, sometimes strongly consistent, but with no answer (Ex. 6.2).



(c) Disintegration either only consistent or strongly consistent, quotient problem either well posed or not (Ex.: 6.3).



(d) A set of uniqueness with no optimal pair and well posed quotient problem (Ex.: 6.4).

Figure 3: The dependence on Γ. In the picture you find, in bold, the set where c is finite.

Three extremal points in $\Pi(\mu, \nu)$ are, for $i \in \{0, 1, -1\}$,

$$\pi_i = \big(\mathrm{Id}, \mathrm{Id} + i\alpha (\mathrm{mod}\ 1)\big)_\sharp \mathcal{L}^1 \quad \Longrightarrow \quad \int c(x,y) d\pi_i = c_i,$$

the optimal one will be the one corresponding to the lowest c_i.

Fix the attention on π_0, which is c-cyclically monotone when $c_1 + c_{-1} \geq 2c_0$. Take as Γ the diagonal $\{x = y\}$: the equivalence classes are given by $\{x + n\alpha \mod 1\}$, the quotient is a Vitali set, and thus the unique consistent disintegration is the trivial disintegration

$$\mathcal{L}^1 = \int \mathcal{L}^1 m(dt).$$

Moreover, one can verify that there is no choice of Γ for which the disintegration is strongly consistent. When $c_1 < c_0$, we have a c-cyclically monotone transference plan π for which the decomposition gives a disintegration which is not strongly consistent and π is not optimal. When $c_1, c_2 > c_0$, we have an optimal c-cyclically monotone transference plan π for which the disintegration consistent with the decomposition in cycles is not strongly consistent.

Example 6.2. Consider an example given in [2], page 135: $\mu = \nu = \mathcal{L}^1$ with the cost given by

$$c(x,y) = \begin{cases} 1 & y - x = 0 \\ 2 & y - x = \alpha \ (\mathrm{mod}\ 1) \\ +\infty & \text{otherwise} \end{cases} \qquad \text{with } \alpha \in [0,1].$$

The extremal plans in $\Pi(\mu,\nu)$ with finite costs are, for $i \in \{0,1\}$,

$$\pi_i = \big(\mathrm{Id}, \mathrm{Id} + i\alpha (\mathrm{mod}\ 1)\big)_\sharp \mathcal{L}^1 \quad \Longrightarrow \quad \int c(x,y) d\pi_i = 1 + i;$$

both are c-cyclically monotone, and the optimal one is π_0. Take $\Gamma = \{x = y\}$: then there is no cycle of finite cost, therefore the cycle decomposition gives classes consisting in singletons, the quotient space is the original one, $m = \mathcal{L}^1$, $\pi_\alpha = \delta_{\{(x,y)\}}$, where α is the class of (x,y). This means that the measurability condition is satisfied, but the quotient problem (which here is essentially the original one) has not uniqueness. Take instead $\Gamma = \{c < \infty\}$: now we have cycles, all with zero cost, obtained by going on and coming back along the same way; consider for example the cycle

$$(w_1, w_1) = (0,0) \ \to \ (w_2, w_2) = (0,\alpha) \ \to \ (w_3, w_3) = (\alpha,\alpha) \ \to \ (w_4, w_4) = (0,0).$$

The situation is similar to Example 6.1 and, as it was there, the disintegration is not strongly consistent. Thus we have that, depending on Γ, strong consistency can be satisfied or not, and when it is, the quotient problem has not uniqueness. This behavior holds when testing either π_0 or π_1, thus it does not depend on the optimality of the plan we are testing.

Example 6.3. Consider the same setting as in Example 6.2, but put the cost to be finite, say zero, also on the lines $\{x = 1\}$ and $\{y = 1\}$. Now, considering $\pi = (\mathrm{Id}, \mathrm{Id})_\sharp \mathcal{L}^1$,

- with Γ containing $(1,1)$, all the points are connected by a cycle of finite cost, we have just one class and optimality follows by c-cyclical monotonicity;
- with $\Gamma = \{(x,x) : \ x \in [0,1)\}$ the classes are made of single points, the disintegrations is trivially measurable, the quotient problem is essentially the original one and we are in the non-uniqueness case;
- when you consider instead $\Gamma = \{(x,x) : \ x \in [0,1)\} \cup \{(x, x+\alpha) : x \in [0,1] \setminus \{1-\alpha\}\}$ again the quotient space is a Vitali set, the strong consistency of the disintegration is lost.

Depending on the choice of Γ, we can have or not strong consistency; moreover, when we have strong consistency, the quotient space can have uniqueness or non-uniqueness. Notice that since there exists Γ for which Theorem 5.6 holds, π must optimal: the first argument does not hold for $(\mathrm{Id}, \mathrm{Id} + \alpha)_\sharp \mathcal{L}^1$, since $\Gamma = \{(x, x + \alpha \mod 1), x \in [0,1]\}$ is not c-cyclically monotone.

Example 6.4 (A set of uniqueness with nonexistence of ϕ, ψ). Consider $\mu = \nu$ with the cost given by

$$c(x,y) = \begin{cases} 1 & y = x \\ 1 - \sqrt{y-x} & y - x = 2^{-n} \\ +\infty & \text{otherwise} \end{cases} \qquad \text{with } n \in \mathbb{N}.$$

Unless μ is purely atomic with a finite number of atoms, there is no optimal potential. However, applying the procedure one can deduce optimality: $\{c < \infty\}$ is acyclic and therefore the cycle decomposition consists in singletons, the quotient spaces are the original ones, and therefore A' of (5.3) is a set of uniqueness, being contained in $\{x \leq y\}$, and $\pi_{\{(x,x)\}} = \delta_{\{(x,x)\}}$, $m = \mu$.

Example 6.5. The final example shows that in the case of consistency only, then we can construct a cost $\tilde{c}$ such that the image measure m is the same but there are non optimal transference plans. We just sketch the main steps.

Let $h_X, h_Y : [0,1] \to [0,1]$ be the Borel quotient maps for the equivalence relation of Definition A.5.

Step 1. The conditional probabilities μ_α, ν_β cannot be purely atomic for m-a.e. α, β.

By the Borel regularity of the map $\alpha \mapsto \mu_\alpha$, one can in fact show ([5]) that there exists a Borel set B such that $B \cap h_X^{-1}(\alpha)$ is countable and the atomic part of μ_α is concentrated on B. Hence if μ_α is purely atomic we can reduce to the case where $h_X^{-1}(\alpha)$ is countable for all α.

Assume by contradiction that each equivalence class has countably many counterimages. We can use Lusin Theorem (Theorem 5.10.3 in [19]) to find a countable family of Borel maps $h'_n : [0,1] \supset B_n \to [0,1]$, $B_n \in \mathcal{B}([0,1])$, $n \in \mathbb{N}$, such that $h_X \circ h'_n = \mathbb{I}_{\llcorner B_n}$ and

$$\mathrm{graph}(h_X) = \bigcup_n \mathrm{graph}(h'_n), \quad \mathrm{graph}(h'_n) \bigcap \mathrm{graph}(h'_m) = \emptyset.$$

Define the analytic $\bar{E}$-saturated sets ($\bar{E}$ is the closed cycles equivalence relation, Definition 5.4)

$$Z_n = P_1\big(\bar{E} \cap [0,1] \times h'_n(B_n)\big) \setminus \bigcup_{i=1}^{n-1} Z_i.$$

By construction, $h'_n(B_n) \cap Z_n$ is an analytic section of Z_n, so that Proposition A.9 implies that the disintegration is strongly consistent. The same clearly holds for $\cup_n Z_n$.

Step 2. We restrict to the case where μ_α, ν_α have no atoms.

The previous step shows that there is a set of positive m-measure for which the conditional probability μ_α is not purely atomic. Let $\mu_{\alpha,c}$ be the continuous part of μ_α: in [5] it is shown that

$$\int \mu_{\alpha,c} m(d\alpha) = \mu_{\llcorner C}$$

for some Borel set C, so that we can assume C compact and restrict the transport to $C \times [0,1]$.

Repeating the procedure for Y, there exists D compact such that for the transport problem in $C \times D$ the conditional probabilities μ_α, ν_α are continuous.

Step 3. We redefine the cost in the set $C \times D$ is order to have the same equivalence classes for h_X, h_Y but for which there are non optimal cyclically monotone costs.

Define the map

$$H_X(\alpha, x) = \mu_\alpha((0,x)), \quad H_Y(\beta, y) = \nu_\beta((0,y)).$$

By Lemma 4.11, we can assume that H_X, H_Y are Borel in (α, x) and (β, y), respectively. If $\bar{c}$ is the cost of Example 6.2, then define

$$\tilde{c}(x,y) = \begin{cases} \bar{c}(H_X(\alpha,x), H_Y(\alpha,y)) & (x,y) \in (h_X \otimes h_Y)^{-1}(\alpha,\alpha) \\ +\infty & \text{otherwise} \end{cases}$$

With the notation of Example 6.2 for π_0 and π_1, for any pseudoinverse $H_X^{-1}(\alpha)$, $H_Y^{-1}(\alpha)$ it is fairly easy to verify that

$$\pi = \int (H_X^{-1}(\alpha) \otimes H_Y^{-1}(\alpha))_\sharp \pi_1 m(d\alpha)$$

is a $\tilde{c}$-cyclically monotone transference plan which is not optimal: the optimal is

$$\pi' = \int (H_X^{-1}(\alpha) \otimes H_Y^{-1}(\alpha))_\sharp \pi_0 m(d\alpha).$$

6.2. **Analysis of the transport problem in the quotient space.** In this section we consider some examples related to the study of the quotient transport problem. The examples are as follows.

(1) The regularity properties of the original cost (e.g. l.s.c.) are in general not preserved (Example 6.6).
(2) In general, there is no way to construct a quotient cost c independently of the transference plan π and different from $\mathbb{1}_{A'}$ (Example 6.7).
(3) The set $\Pi^f(m,m)$ strictly contains the set $(h_X \otimes h_Y)_\sharp \Pi^f(\mu,\nu)$ (Examples 6.7, 6.8).

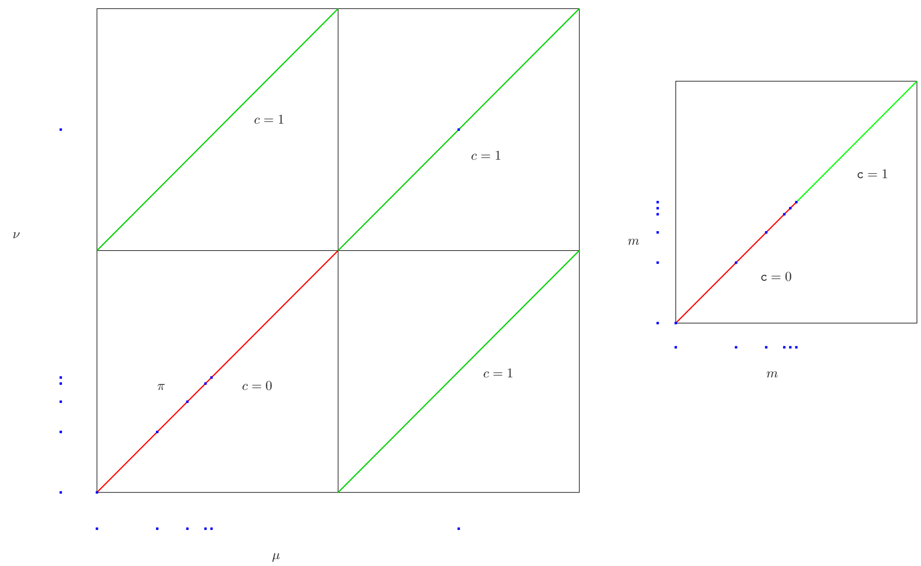


Figure 4: The cost of Example 6.6. Outside the diagonal segments the cost is $+\infty$.

(4) If the uniqueness assumption of Theorem 5.6 does not hold, then we can construct a cost c' which gives the same equivalence classes and quotient transport problem and such that the original π is c'-cyclically monotone but not optimal for c' (Proposition 6.9).

Example 6.6 (Fig. 4). Consider the cost

$$c(x,y) = \begin{cases} 0 & y = x, x \in [0,1/2] \\ 1 & y = x + 1/2 \mod 1 \\ 1 & y = x, x \in (1/2,1) \end{cases}$$

and the measures

$$\mu = \nu = \sum_{i=1}^{+\infty} 2^{-i-1} \delta\bigg(x - \frac{1}{2} + 2^{-i}\bigg) + \frac{1}{2}\delta(x - 3/2), \qquad \pi = (x,x)\sharp\mu.$$

If we require that in each equivalence class

$$\mathsf{c}_\pi(\alpha,\alpha) = \int c(x,y)\pi_\alpha(dxdy), \qquad \pi_\alpha \in \Pi(\mu_\alpha,\nu_\alpha), \tag{6.1}$$

one obtains

$$\mathsf{c}(\alpha,\beta) = \begin{cases} 0 & \beta = \alpha = 1/2 - 2^{-i}, i \in \mathbb{N} \\ 1 & \beta = \alpha = 1/2 \end{cases}$$

Clearly this cost is not l.s.c., and there is no way to make it l.s.c. under (6.1). This example shows that we cannot preserve regularity properties for the quotient cost c.

Example 6.7 (Fig. 5). Let $r \in [0,1/4] \setminus \mathbb{Q}$ and consider the cost

$$c(x,y) = \begin{cases} 1 & x = y, x \in [0,1) \\ 1+d & y = x + 1/2, x \in [0,1/2) \\ 1+d & y = x - 1/2, x \in [1/2,1) \\ 0 & y = x + r, x \in [0,1/2 - r) \\ e & y = x + r, x \in [1/2, 1-r) \\ 0 & y = x - 1/2 + r, x \in [1-r,1) \\ +\infty & \text{otherwise} \end{cases} \qquad d, e \geq 0.$$

The settings are

$$\mu = \nu = \mathcal{L}^1, \qquad \Gamma = \{y = x\}.$$

The equivalence relation is $(x,x)E(x+1/2, x+1/2)$: for simplicity we consider the quotient space as $[0,1/2)$.

In the quotient space, the cost c_π is finite only on $y = x$ and $y = x + r \mod 1/2$. We now consider two particular transference plans.

The easiest one is $\pi_0 = (x, f_0(x))_\sharp \mathcal{L}^1$, where

$$f_0(x) = \begin{cases} x & x \in [0, \frac{1}{2}) \\ x + r & x \in [\frac{1}{2}, 1 - r) \\ x - \frac{1}{2} + r & x \in [1 - r, 1) \end{cases}$$

for which by formula (6.1) we obtain a quotient cost of

$$\mathsf{c}_0 = \begin{cases} 1 & \beta = \alpha, \alpha \in [0, 1/2) \\ e & \beta = \alpha + r, \alpha \in [0, \frac{1}{2} - r) \\ 0 & \beta = \alpha - \frac{1}{2} + r, \alpha \in [\frac{1}{2} - r, \frac{1}{2}) \end{cases} \tag{6.2}$$

Another cost is obtained by $\pi_1 = (x, f_1(x))_\sharp \mathcal{L}^1$, where

$$f_1(x) = \begin{cases} x + r & x \in [0, \frac{1}{2} - r) \\ x + \frac{1}{2} & x \in [\frac{1}{2} - r, \frac{1}{2}) \\ x - \frac{1}{2} & x \in [\frac{1}{2}, \frac{1}{2} + r) \\ x & x \in [\frac{1}{2} + r, 1 - r) \\ x - \frac{1}{2} + r & x \in [1 - r, 1) \end{cases}$$

In this case the cost is by (6.1)

$$\mathsf{c}_1 = \begin{cases} 1 + d & \beta = \alpha, \alpha \in [0, r) \cup [1/2 - r, 1/2) \\ 1 & \beta = \alpha, \alpha \in [r, 1/2 - r) \\ 0 & \beta = \alpha + r \mod 1/2 \end{cases} \tag{6.3}$$

Since it is impossible to have a transference plan π in the original coordinates such that

$$\mathsf{c}_\pi = \begin{cases} 1 & \beta = \alpha, \alpha \in [0, 1/2) \\ 0 & \beta = \alpha + r \mod 1/2 \end{cases}$$

then it follows that there is no clear way to associate the cost c in the quotient space independently of the transport plan π.

We note that there is no transference plan whose image is concentrated only on $\beta = \alpha + r \mod 1$, so that in general the image of $\Pi^f(\mu, \nu)$ under the map $(h_X \otimes h_Y)$ is a strict subset of $\Pi^f(m, m)$.

Example 6.8 (Fig. 6). We consider the cost for $r \in [\frac{1}{4}, \frac{1}{2}] \setminus \mathbb{Q}$

$$c(x,y) = \begin{cases} 1 & y = x, x \in [0,1) \\ 1 + d & y = \frac{x}{2} + \frac{1}{2}, x \in [0,1) \\ 1 + d & y = 2x - 1, x \in [0,1) \\ e & y = x + r, x \in [0, \frac{1}{2} - r) \\ f & y = x - 2^{-i}(\frac{1}{2} - r), x \in (1 - 2^{-i}) + 2^{-i}[\frac{1}{2} - 2^{-i+1}r, \frac{1}{2} - 2^{-i}r), i \in \mathbb{N} \\ +\infty & \text{otherwise} \end{cases}.$$

We consider the measures

$$\mu = \nu = \frac{3}{2} \sum_{i=0}^{+\infty} 2^{-i} \mathcal{L}^1 \llcorner_{[1-2^{-i}, 1-2^{-i-1})}.$$

Since the measure of the segment $[1 - 2^{-i}, 1 - 2^{-i-1})$ is 2^{-2i-1}, all measures π with finite cost in $\Pi(\mu, \nu)$ are concentrated on the segments

$$\big\{y = x, x \in [0,1]\big\} \cup \big\{y = x/2 + 1/2, x \in [0,1]\big\} \cup \big\{y = 2x - 1, x \in [1/2, 1]\big\}.$$

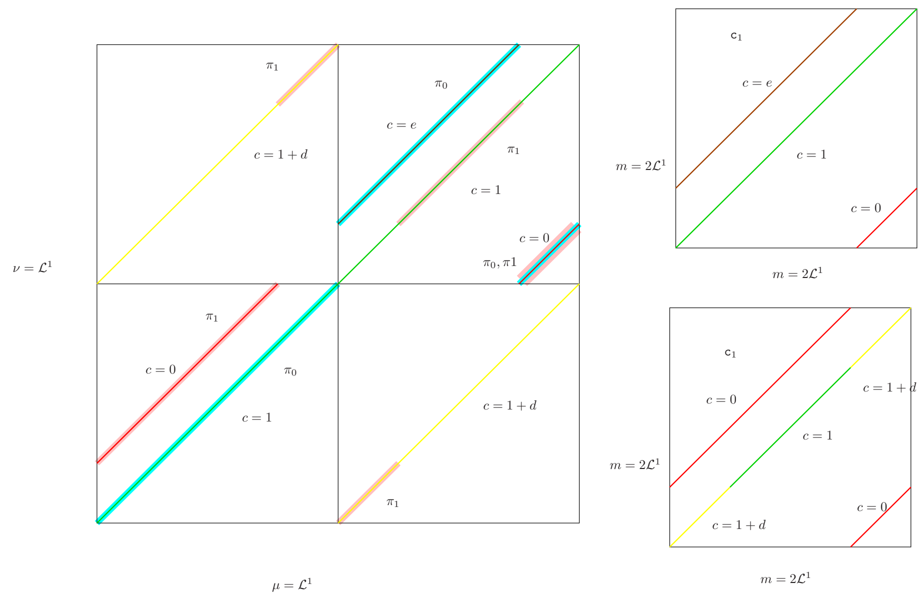


Figure 5: The cost of Example 6.7. Outside the segments the cost is $+\infty$, while for the two different tranference plans the quotient costs are given by (6.2), (6.3).

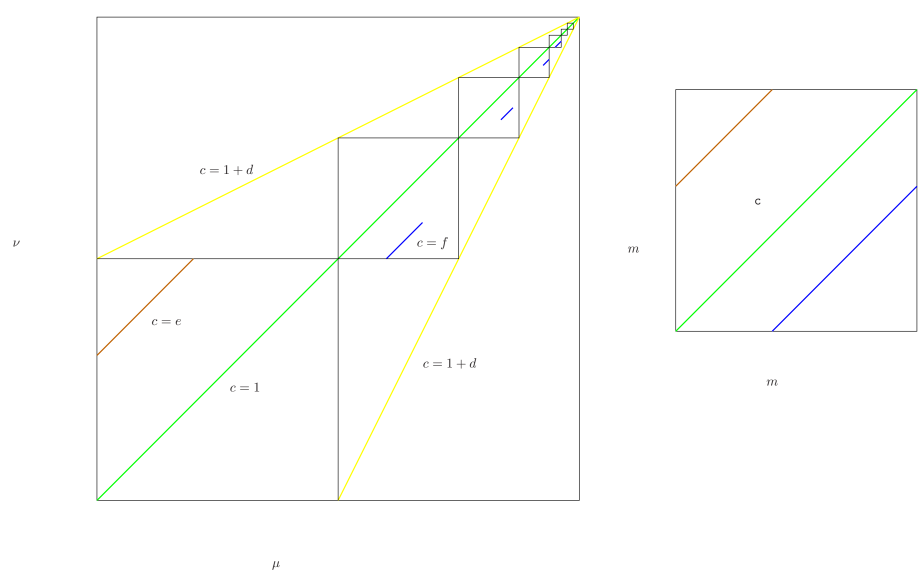


Figure 6: The cost of Example 6.8.

This can be seen in the quotient space, because

$$m = \mathcal{L}^1,$$

and every measure $\tilde{m} \in \Pi^f(m, m)$ is of the form $\tilde{m} = a_1(x, x)_\sharp \mathcal{L}^1 + a_2(x, x + r \bmod 1)_\sharp \mathcal{L}^1$, $a_1, a_2 \geq 0$ and $a_1 + a_2 = 1$. But clearly this cannot be any image of a measure with finite cost in $\Pi^f(\mu, \nu)$.

The next proposition shows that if A' is not a set of uniqueness, then the problem of optimality cannot be decided by just using c-cyclical monotonicity.

Proposition 6.9. *If there exists a transference plan $\tilde{m} \in \Pi^f(m,m)$ different from $(\mathbb{I},\mathbb{I})_\sharp m$, then there exists a cost $\hat{c}(x,y)$ for which the following holds:*

(1) *the set Γ is $\hat{c}$-cyclically monotone;*
(2) *there are two measures π, $\tilde{\pi}$ in $\Pi(\mu,\nu)$ such that*

$$\pi(\Gamma) = 1, \qquad \int \hat{c}\tilde{\pi} < \int \hat{c}\pi < +\infty.$$

A variation of the following proof (using Lusin Theorem and inner regularity) allows to construct a cost which is also l.s.c. if the original cost is.

Proof. Let $\tilde{m} \in \Pi^f(m,m) \setminus \{(\mathbb{I},\mathbb{I})_\sharp m\}$, and consider a Borel cost c such that

$$\mathsf{c}\big([0,1]^2 \setminus (h_X \otimes h_Y)(\{c < +\infty\})\big) = +\infty, \qquad \int \mathsf{c}\tilde{m} < \int \mathsf{c}m < +\infty.$$

It is fairly easy to construct such a cost.

Define now

$$\Gamma := (h_X \otimes h_Y)^{-1}(\{\alpha = \beta\}), \quad \hat{c}(x,y) := \mathsf{c}(h_X(x), h_Y(y)),$$
$$\pi := \int \mu_\alpha \otimes \nu_\alpha m(d\alpha), \quad \tilde{\pi} := \int \mu_\alpha \otimes \nu_\beta \tilde{m}(d\alpha d\beta).$$

It follows that

$$\int \hat{c}\tilde{\pi} = \int \mathsf{c}\tilde{m} < \int \mathsf{c}m = \int \hat{c}\pi < +\infty.$$

Moreover, since $h_X \otimes h_Y(\{c < +\infty\})$ is acyclic w.r.t. $\{\alpha = \beta\}$ (by the same proof of Lemma 4.7), the equivalence classes w.r.t. the closed cycles equivalence relation $\bar{E}$ do not change, so that Γ is acyclic. □

7. Existence of an optimal potential

Let $\pi \in \mathcal{P}([0,1]^2)$ be concentrated on a c-cyclically monotone set Γ. Assume that there exist partitions $\{X_\alpha\}_\alpha$, $\{Y_\beta\}_\beta$ of $[0,1]$ into Borel sets such that

- $\Gamma \subset \cup_\alpha X_\alpha \times Y_\alpha$ — i.e. Γ satisfies the crosswise condition of Definition 2.3 w.r.t. the partition;
- in each set $X_\alpha \times Y_\alpha$ there exist Borel optimal potentials ϕ_α, ψ_α:

$$\phi_\alpha + \psi_\alpha \leq c \text{ on } X_\alpha \times Y_\alpha \qquad \phi_\alpha + \psi_\alpha = c \text{ on } \Gamma \cap X_\alpha \times Y_\alpha.$$

Is it possible to find a Borel couple of functions ϕ, ψ s.t.

$$\phi + \psi \leq c \text{ on } \cup_\alpha X_\alpha \times Y_\alpha \qquad \phi + \psi = c \ \pi\text{-a.e.?}$$

We show that this is the case under Assumption 1, i.e. if the disintegration of π w.r.t the partition $\{X_\alpha \times Y_\alpha\}$ is strongly consistent. If $\{c < +\infty\} \subset \cup_\alpha X_\alpha \times Y_\alpha$ this provides clearly an optimal couple.

The approach is to show that the set

$$\Big\{(\alpha, \tilde{\phi}, \tilde{\psi}) : \ \tilde{\phi}, \tilde{\psi} \text{ optimal couple in } X_\alpha \times Y_\alpha\Big\}$$

is an analytic subset of a suitable Polish space, that we are first going to define. We then apply a selection theorem to construct an optimal couple.

In order to structure the ambient space with a Polish topology, we need some preliminary lemmas.

Lemma 7.1. *For every nonnegative function $\bar{\varphi} \in C^0([0,1])$ the map*

$$G_{\bar{\varphi}} \quad : \quad \mathcal{M}([0,1]) \quad \ni \quad \mu \quad \mapsto \quad \int \bar{\varphi}\mu^+ \quad \in \quad \mathbb{R}$$

is convex l.s.c. is w.r.t. weak-topology.*

Proof. Since for every $\mu \in \mathcal{M}([0,1])$

$$\sup\bigg\{ \int \varphi\mu : \ 0 \leq \varphi \leq \bar{\varphi} \bigg\} = \int \bar{\varphi}\mu^+,$$

then $G_{\bar{\varphi}}$ is the supremum of bounded linear functionals, proving the thesis. □

Corollary 7.2. *The map*

$$\mathcal{M}([0,1]) \quad \ni \quad \mu \quad \mapsto \quad \mu^+ \quad \in \quad \mathcal{M}^+([0,1])$$

is Borel w.r.t. weak-topology. For every nonnegative measure ξ the sublevel set $\{\mu : \ \mu^+ \leq \xi\}$ is closed and convex: in fact $\mu \mapsto \mu^+$ is order convex, meaning that*

$$(\lambda\mu + (1-\lambda)\nu)^+ \leq \lambda\mu^+ + (1-\lambda)\nu^+.$$

Proof. It is enough to observe that any function $f : \mathcal{M}([0,1]) \to \mathcal{M}([0,1])$ is Borel if and only if the function $\mu \mapsto \int \varphi f(\mu)$ is Borel for every nonnegative $\varphi \in C^0([0,1])$: the Borel measurability then follows by Lemma 7.1. As well, f is order convex if and only if $\mu \mapsto \int \varphi f(\mu)$ is convex $\forall \varphi \in C^0([0,1];\mathbb{R}^+)$. □

Corollary 7.3. *The function*

$$\mathcal{M}([0,1]) \times \mathcal{M}([0,1]) \quad \ni \quad (\mu_1,\mu_2) \quad \mapsto \quad \mu_1 \wedge \mu_2 \quad \in \quad \mathcal{M}([0,1])$$

is Borel w.r.t. weak-topology.*

Proof. The thesis follows by the relation $\mu_1 \wedge \mu_2 = \mu_1 - [\mu_1 - \mu_2]^+$ and Corollary 7.2. □

Lemma 7.4. *The function*

$$\mathcal{M}^+([0,1])^3 \times C^0([0,1];\mathbb{R}^+) \ni (\mu_1,\mu_2,\mu_3,\phi) \ \mapsto \ \int \phi \frac{d\mu_2}{d\mu_1}\frac{d\mu_3}{d\mu_1}\mu_1 \in [0,+\infty]$$

is Borel w.r.t. weak-topology.*

Proof. Let $\{h_{j,I}\}_{j=1}^{2^I}$ be a partition of $[0,1]$ into continuous functions such that

$$0 \leq h_{j,I} \leq 1, \ \sum_{j=1}^{2^I} h_{j,I} = 1, \ \operatorname{supp} h_{j,I} \subset \big[(j-1)2^{-I} - 2^{-I-2}, j2^{-I} + 2^{-I-2}\big].$$

Define the l.s.c. and continuous functions, respectively,

$$\mathbb{R}^+ \ni x \mapsto x^{-1^*} := \begin{cases} 0 & x = 0 \\ 1/x & x > 0 \end{cases}$$

$$\mathcal{M}([0,1]) \times C^0([0,1]) \ni (\mu,\phi) \mapsto \bigg(\int h_{j,I}\phi\mu \bigg)_{j=1}^{2^I} \in \mathbb{R}^{2^I}$$

If $\mu_2 = (d\mu_2/d\mu_1)\mu_1$, then

$$g_I(\mu_1,\mu_2) := \sum_{j=1}^{2^I} h_{j,I}(x) \bigg(\int h_{j,I}\mu_1 \bigg)^{-1^*} \bigg(\int h_{j,I}\mu_2 \bigg) \ \to \ \frac{d\mu_2}{d\mu_1}$$

in $L^1(\mu_)$: in fact, for continuous functions the resut follows by uniform continuity, and for the general case one observes that

$$\int \bigg| \sum_{j=1}^{2^I} h_{j,I}(x) \bigg(\int h_{j,I}\mu_1 \bigg)^{-1^*} \bigg(\int h_{j,I} f \mu_1 \bigg) \bigg| \mu_1(dx) \leq \sum_{j=1}^{2^I} \bigg| \int h_{j,I} f \mu_1 \bigg| \leq \|f\|_{L^1(\mu_1)}.$$

For $\phi \in C^0([0,1])$ and $0 \leq \mu_2,\mu_3 \leq k\mu_1$ it follows

$$\int \phi \frac{d\mu_2}{d\mu_1}\frac{d\mu_3}{\mu_1}\mu_1 = \lim_{I\to+\infty} \int \phi g_I(\mu_1,\mu_2) g_I(\mu_1,\mu_3)\mu_1.$$

We finally reduce to the case $\mu_2 \leq k\mu_1$ and $\mu_3 \leq k\mu_1$: indeed

$$\int \phi \frac{d\mu_2}{d\mu_1}\frac{d\mu_3}{\mu_1}\mu_1 = \lim_{k\to\infty}\lim_{I\to+\infty} \int \phi g_I(\mu_1,\mu_2 \wedge (k\mu_1)) g_I(\mu_1,\mu_3 \wedge (k\mu_1))\mu_1$$

and by Corollary 7.3 the map

$$(\mu_1,\mu_2,\mu_3) \mapsto \Big(\mu_1, (k\mu_1) \wedge \mu_2, (k\mu_1) \wedge \mu_3\Big)$$

is Borel. By composition of the above Borel maps, the statement of the lemma is proved. □

Lemma 7.5. *The function*

$$\begin{array}{rcccc} H_M & : & \mathcal{M}^+([0,1])^2 \times \mathcal{M}([0,1])^2 \times \mathcal{M}^+([0,1]^2) & \to & [0,+\infty] \\ & & (\mu,\nu,\eta,\xi,\pi) & \mapsto & H_M := \int \Big(\frac{d(\eta+M\mu)^+}{d\mu}\Big)\Big(\frac{d(P_1)_\sharp\pi}{d\mu}\Big)\mu \\ & & & & + \int \Big(\frac{d(\xi+M\nu)^+}{d\nu}\Big)\Big(\frac{d(P_2)_\sharp\pi}{d\nu}\Big)\nu \end{array}$$

is Borel w.r.t. weak*-topology for all $M \in \mathbb{R}^+$.

Proof. It follows immediately from Corollary 7.2 and Lemma 7.4. □

Lemma 7.6. *The subset of sequences in $\mathbb{R}^{\mathbb{N}}$ converging to zero is analytic w.r.t. the product topology.*

Proof. The family on nondecreasing sequences m_n is a closed subset of $\mathbb{N}^{\mathbb{N}}$, with the product topology. The sequences of $\mathbb{R}^{\mathbb{N}}$ converging to zero are then the projection of the subset of $\mathbb{R}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}}$

$$C := \Big\{ (\{f_\ell\}_{\ell\in\mathbb{N}}, \{m_n\}_{n\in\mathbb{N}}) : \ |f_i| \leq 2^{-n} \ \forall i \geq m_n \Big\}.$$

We now show that C is closed in the product topology, from which the result follows.
Consider sequences $\{f_{\ell,k}\}_\ell$, $\{m_{n,k}\}_n$ converging pointwise to $\{f_\ell\}_\ell$, $\{m_n\}_n$, with $(\{f_{\ell,k}\}_\ell, \{m_{n,k}\}_n) \in C$. Then for each $n \in \mathbb{N}$ there exists $k(n)$ such that the sequence $\{m_{n,k}\}_k$ is constantly m_n for $k > k(n)$. As a consequence, for all $k > k(n)$ one has $|f_{i,k}| \leq 2^{-n}$ for $i \geq m_n$. Since $\{f_{i,k}\}_i$ converges pointwise, it follows that $|f_i| \leq 2^{-n}$ for $i \geq m_n$. Hence $(\{f_\ell\}_\ell, \{m_n\}_n) \in C$. □

Given a subset J of $\mathbb{R} \cup \{\pm\infty\}$, we denote by $L(\mu; J)$ the μ-measurable maps from $[0,1]$ to J. If not differently stated, μ-measurable functions are equivalence classes of functions which coincide μ-a.e..

Proposition 7.7. *There exists a Polish topology on linear space*

$$L = \Big\{ (\mu,\varphi) : \ \mu \in \mathcal{P}([0,1]), \ \varphi \in L(\mu; \mathbb{R}\cup\{\pm\infty\}) \Big\}$$

such that the map

$$\begin{array}{rcccc} I & : & L & \to & \mathcal{P}([0,1]) \times \prod_{M=1}^\infty \mathcal{M}([0,1]) \\ & & (\mu,\varphi) & \mapsto & \big(\mu, \big\{(\varphi\wedge M)\vee(-M)\mu\big\}_{M\in\mathbb{N}}\big) \end{array}$$

is continuous.

Proof. We inject L in $\mathcal{P}([0,1]) \times \prod_{M=1}^\infty \mathcal{M}([0,1])$ by the map I. The image of L is the set

$$\mathrm{Im}(I) = \Big\{ (\mu, \eta_M) : \ \eta_N = (\eta_M \wedge N\mu) \vee (-N\mu) \quad \text{for } M > N \Big\}. \tag{7.1}$$

Notice that the compatibility condition $\eta_N = (\eta_M \wedge N\mu) \vee (-N\mu)$ implies that the Radon-Nikodym derivative $\varphi_N := \frac{d\eta_N}{d\mu}$ converges μ-a.e. to a uniquely identified $\varphi \in L(\mu; \mathbb{R}\cup\{\pm\infty\})$.

We observe that by Corollary 7.3 the function

$$(\mu,\eta) \mapsto F_N(\mu,\eta) := -(-(\eta \wedge N\mu) \wedge N\mu)$$

is Borel and $\mathrm{Im}(I)$ is the intersection of the following countably many graphs

$$\mathrm{Im}(I) = \bigcap_{N<M} \Big\{ (\mu, \{\eta_Q\}_Q) : F_N(\mu,\eta_M) = \eta_N \Big\}.$$

Being a Borel subset of a Polish space, by Theorem 3.2.4 of [19] there is a finer Polish topology on $\mathcal{P}([0,1]) \times \prod_{M=1}^\infty \mathcal{M}([0,1])$ such that $\mathrm{Im}(I)$ itself is Polish, and this Polish topology can be pulled back to L by the injective map I. The continuity of I, also w.r.t. the product weak* topology on the image space, is then immediate. □

Lemma 7.8. *The subset $L^f := \{(\mu,\varphi) : \ \varphi \in L(\mu;\mathbb{R})\}$ of L is analytic.*

Proof. If $\varphi \in L(\mu; \mathbb{R}\cup\{\pm\infty\})$, the condition $\varphi \in L(\mu;\mathbb{R})$ is clearly equivalent to $\lim_M \|\mu\|(|\varphi| > M) = 0$. Since the injection I is continuous and $I(L)$ is Borel by Proposition 7.7, it is enough to prove that

$$I(L^f) = \Big\{ (\mu, \{\xi_M\}_M) : \ \lim_{M\to+\infty} \|\xi_{M+1} - \xi_M\| = 0 \Big\}$$

is analytic. By Lemma 7.6 this follows by the l.s.c. of the map

$$\begin{array}{ccc} \mathcal{P}([0,1]) \times \prod_{M=1}^{\infty} \mathcal{M}([0,1]) & \to & \mathbb{R}^{\mathbb{N}} \\ (\mu, \{\xi_M\}_M) & \mapsto & \big\{ \|\xi_{M+1} - \xi_M\| \big\}_M \end{array}.$$

□

Theorem 7.9. *Let c be l.s.c.. Assume that the disintegration of π w.r.t. a partition $\{X_\alpha \times Y_\alpha\}_\alpha$ is strongly consistent and that there exist optimal potentials $\phi_\alpha \in \mathcal{B}(X_\alpha, \mathbb{R} \cup \{-\infty\})$, $\psi_\alpha \in \mathcal{B}(Y_\alpha; \mathbb{R} \cup \{-\infty\})$:*

$$\phi_\alpha + \psi_\alpha \leq c \text{ on } X_\alpha \times Y_\alpha \qquad \phi_\alpha + \psi_\alpha = c \text{ on } \Gamma \cap X_\alpha \times Y_\alpha.$$

Then there exist Borel optimal potentials on $\cup_\alpha X_\alpha \times Y_\alpha$.

Proof. We prove the theorem by means of Von Neumann's selection principle.

Step 1. Consider the Polish space

$$Z := L \times L \times \mathcal{P}([0,1]^2).$$

We first prove the analyticity of the set $A \subset Z$ made of those

$$((\mu, \varphi), (\nu, \psi), \pi) \in L^f \times L^f \times \mathcal{P}([0,1]^2)$$

satisfying the relations

(1) $(P_1)_\sharp \pi = \mu$, $(P_2)_\sharp \pi = \nu$;
(2) $\phi + \psi \leq c$ out of cross-negligible sets w.r.t. the measures μ, ν;
(3) $\phi + \psi = c$ π-a.e..

Since Σ^1_1 is closed under countable intersections, it suffices to show that each of the conditions above defines an analytic set.

Constraint (1) defines a closed set, by the continuity of the immersion I in Proposition 7.7 and because $\{(\mu, \nu, \pi) : \pi \in \Pi(\mu, \nu)\}$ is compact in $\mathcal{P}([0,1]) \times \mathcal{P}([0,1]) \times \mathcal{P}([0,1]^2)$.

Setting $\phi_M = ((\phi \wedge M) \vee (-M))$, $\psi_M = ((\psi \wedge M) \vee (-M))$ for $M \in \mathbb{N}$, Condition (2) is equivalent to

$$\int \phi_M (P_1)_\sharp \pi + \int \psi_M (P_2)_\sharp \pi \leq \int c\pi \qquad \forall \pi \in \Pi^{\leq}(\mu, \nu), \forall M \in \mathbb{N}. \tag{7.2}$$

Indeed, suppose that Condition (2) is not satisfied, i.e. the set $\{(x, y) : \phi(x) + \psi(y) > c(x, y)\}$ is not cross-negligible. Then, since ϕ_M, ψ_M converge pointwise to ϕ, ψ, the set $\{(x, y) : \phi_M(x) + \psi_M(y) > c(x, y)\}$ cannot be cross-negligible. By the Duality Theorem B.2 there exists a non-zero $\pi \in \Pi^{\leq}(\mu, \nu)$ concentrated on $\{(x, y) : \phi_M(x) + \psi_M(y) > c(x, y)\}$ and therefore (7.2) does not hold. The converse is immediate, as $\phi_M + \psi_M \leq c$.

We consider the Borel set (Lemma 7.5)

$$C_{n,M} := \bigg\{ (\mu, \nu, \xi, \eta, \pi) : \ H_M(\mu, \nu, \xi, \eta, \pi) - \int (c + 2M)\pi \geq 2^{-n}, \pi \in \Pi^{\leq}(\mu, \nu) \bigg\}. \tag{7.3}$$

Since for $\pi \in \Pi^{\leq}(\mu, \nu)$ one has

$$H_M(\mu, \nu, \xi, \eta, \pi) = \int \frac{d(\xi + M\mu)^+}{d\mu} (P_1)_\sharp \pi + \int \frac{d(\eta + M\nu)^+}{d\nu} (P_2)_\sharp \pi$$

then for fixed (μ, ν, η, ξ) the function

$$\pi \mapsto \begin{cases} H_M(\mu, \nu, \xi, \eta, \pi) - \int (c + 2M)\pi & \pi \in \Pi^{\leq}(\mu, \nu) \\ -\infty & \text{otherwise} \end{cases}$$

is u.s.c. for l.s.c. cost c: we have used the fact that

$$\big\{ m \in \mathcal{M}([0,1]) : 0 \leq m \leq \mu \big\} \ni m \mapsto \int f m \in \mathbb{R}$$

is continuous for all $f \in L^1(\mu)$ and $\{0 \leq m \leq \mu\}$ is closed. In particular the section

$$C_{n,M} \cap \{(\mu, \nu, \xi, \eta)\} \times \mathcal{P}([0,1]^2)$$

is closed, hence compact. By Novikov Theorem (Theorem 4.7.11 of [19]), it follows that

$$P_{1234}(C_{n,M}) = \left\{ (\mu,\nu,\eta,\xi) : \exists \pi \in \Pi^{\leq}(\mu,\nu), H_M(\mu,\nu,\eta,\xi,\pi) - \int (c+2M)\pi \geq 2^{-n} \right\}$$

is Borel. Finally, the set

$$D_M := \bigcup_{n\in\mathbb{N}} P_{1234}(C_{n,M}) = \left\{ (\mu,\nu,\eta,\xi) : \exists \pi \in \Pi^{\leq}(\mu,\nu), H_M(\mu,\nu,\eta,\xi,\pi) - \int (c+2M)\pi > 0 \right\}$$

is Borel.

Condition (7.2) thus can be rewritten as

$$\left\{ \big(\mu,\nu,\{\xi_M\}_M,\{\eta_M\}\big) \in \mathcal{P}([0,1])^2 \times \left(\prod_{M=1}^{\infty} \mathcal{M}([0,1]) \right)^2 : (\mu,\nu,\xi_M,\eta_M) \notin D_M \right\},$$

and the above discussion implies that this is a Borel set.

We prove finally that Condition (3) identifies an analytic set. Consider the map

$$\left(\prod_{M=1}^{\infty} \mathcal{M}([0,1]) \times \mathcal{M}([0,1]) \right) \times \mathcal{P}([0,1]^2) \ni (\{\xi_M,\eta_M\}_M,\pi) \mapsto \left\{ \int \xi_M + \int \eta_M - \int c\pi \right\}_M \in \mathbb{R}^{\mathbb{N}}.$$

This function is clearly Borel. Moreover, by Lemma 7.6 the family of sequences converging to 0 is an analytic subset of $\mathbb{R}^{\mathbb{N}}$, and therefore his counterimage is analytic. The thesis follows again by the continuity of the immersion I of Proposition 7.7.

Step 2. Since the set A of Step 1 is analytic, and the map $[0,1] \ni \alpha \mapsto \pi_\alpha \in \mathcal{P}([0,1]^2)$ can be assumed to be Borel, then the set $B = [0,1] \times A \cap \{(\alpha,(\mu,\varphi),(\nu,\psi),\pi : \ \pi = \pi_\alpha)\}$ is analytic.

Step 3. By Von Neumann's selection principle applied to B, there exists an analytic map

$$[0,1] \quad \ni \quad \alpha \quad \mapsto \quad \big((\mu_\alpha,\phi_\alpha),(\nu_\alpha,\psi_\alpha)\big) \quad \in \quad L \times L \ .$$

Hence, by the immersion of I of Proposition 7.7 we can define the sequence of measures

$$\xi_M := \int \xi_{M,\alpha} m(d\alpha) \qquad \eta_M := \int \eta_{M,\alpha} m(d\alpha).$$

It is not difficult to show that $\big(\mu,\{\xi_M\}_{M\in\mathbb{N}}\big)$ and $\big(\nu,\{\eta_M\}_{M\in\mathbb{N}}\big)$ belong to the image (7.1) of I: by the formula

$$\xi_M = \frac{d\xi_M}{d\mu}\mu = \int \left(\frac{d\xi_M}{d\mu}\mu_\alpha \right) m(d\alpha) = \int \xi_{M,\alpha} m(d\alpha)$$

it follows that $(\mu,\{\xi_M\}_M) \in L^f$ and satisfies the compatibility condition. Therefore taking the counterimage with I one can define functions $\phi \in L(\mu)$, $\psi \in L(\nu)$ which are optimal potentials. □

Remark 7.10. Theorem 7.9 does not provide an optimal couple for a generic equivalence relation different from the axial one, and in particular it does not apply for the cycle equivalence relation (see Example 6.4).

Remark 7.11. Even if every two points are connected by an axial path and there exist Borel potentials, in general there is no point $(\bar{x},\bar{y})$ such that the extensions of Corollary C.7 define Borel potentials $\tilde{\phi}$, $\tilde{\psi}$.

Remark 7.12. In the proof one can observe that we can replace the cost c with any other cost c', just requiring that for m-a.e. α it holds $\phi_\alpha + \psi_\alpha \leq c'$. In particular, we can take a cost whose graph is σ-compact in each equivalence class and prove that the sets $C_{n,M}$ of (7.3) are σ-compact.

This shows how Theorem 7.9 can be extended to π-measurable costs.

Appendix A. Disintegration theorem in countably generated probability spaces

In this section we prove the Disintegration Theorem for measures in countably generated σ-algebras, with some applications. The results of this sections can be deduced from Section 452 of [10]; for completeness we give here self-contained proofs.

We consider the following objects:

(1) (X, Ω, μ) a countably generated probability space;
(2) $X = \cup_{\alpha \in \mathtt{A}} X_\alpha$ a partition of X;
(3) $\mathtt{A} = X/\sim$ the quotient space, where $x_1 \sim x_2$ if and only if there exists α such that $x_1, x_2 \in X_\alpha$;
(4) $h : X \to \mathtt{A}$ the quotient map $h(x) = x^\bullet = \{\alpha : x \in X_\alpha\}$.

We can give to $\mathtt{A}$ the structure of probability space as follows:

(1) define the σ-algebra $\mathscr{A} = h_\sharp(\Omega)$ on $\mathtt{A}$ as the *quotient σ-algebra*

$$A \in \mathscr{A} \quad \iff \quad \bigcup_{\alpha \in A} \{x : h(x) = \alpha\} = h^{-1}(A) \in \Omega;$$

(2) define the probability measure $m = h_\sharp \mu$.

We can rephrase (1) by saying that $\mathscr{A}$ is the largest σ-algebra such that $h : X \to \mathtt{A}$ is measurable: it can be considered as the subalgebra of Ω made of all saturated measurable sets.

Definition A.1. The σ-algebra $\mathscr{A}$ is *essentially countably generated* if there is a countable family of sets $A_n \in \mathscr{A}$, $n \in \mathbb{N}$, such that for all $A \in \mathscr{A}$ there exists $\hat{A} \in \mathfrak{A}$, where $\mathfrak{A}$ is the σ-algebra generated by A_n, $n \in \mathbb{N}$, which satisfies $m(A \vartriangle \hat{A}) = 0$.

The first result of this section is the structure of $\mathscr{A}$ as a σ-algebra.

Proposition A.2. *The σ-algebra $\mathscr{A}$ is essentially countably generated.*

Notice that we cannot say that the σ-algebra $\mathscr{A}$ is countably generated: for example, take $([0,1], \mathcal{B})$ and $x^\bullet = \{x + \mathbb{Q}\} \cap [0,1]$. We are stating that the measure algebra $\mathscr{A}/\mathscr{N}_m$, where $\mathscr{N}_m$ is the σ-ideal of m-negligible sets, is countably generated.

The proposition is a consequence of Maharam Theorem, a deep result in measure theory, and can be found in [9], Proposition 332T(b). We give a direct proof of Proposition A.2. The fundamental observation is the following lemma.

Lemma A.3. *Let f_n be a countable sequence of measurable functions on $\mathtt{A}$. Then there is a countably generated σ-subalgebra $\mathfrak{A}$ of $\mathscr{A}$ such that each f_n is measurable.*

Proof. The proof is elementary, since this σ-algebra is generated by the countable family of sets

$$\Big\{ f_n^{-1}(q_m, +\infty), q_m \in \mathbb{Q}, m \in \mathbb{N} \Big\}.$$

This is actually the smallest σ-algebra such that all f_n are measurable. □

Proof of Proposition A.2. The proof will be given in 3 steps.

Step 1. Define the map $\Omega \ni B \to f_B \in L^\infty(m)$ by

$$h_\sharp \mu \llcorner_B = f_B m. \tag{A.1}$$

The map is well defined by Radon-Nikodym theorem, and $0 \leq f_B \leq 1$ m-a.e..

Given an increasing sequence of $B_i \in \Omega$, then

$$\int_A f_{\cup_i B_i} m = \mu(h^{-1}(A) \cap \cup_i B_i) = \lim_i \mu(h^{-1}(A) \cap B_i) = \lim_i \int_A f_{B_i} m = \int_A \lim_i f_{B_i} m,$$

where we have used twice the Monotone Convergence Theorem and the fact that f_{B_i} is increasing m-a.e.. Hence $f_{\cup_i B_i} = \lim_i f_{B_i}$. By repeating the same argument and using the fact that m is a probability measure, the same formula holds for decreasing sequences of sets, and for disjoint sets one obtains in the same way $f_{\cup_i B_i} = \sum_i f_{B_i}$.

Step 2. Let $\mathscr{B} = \{B_n, n \in \mathbb{N}\}$, be a countable family of sets generating Ω: without any loss of generality, we can assume that $\mathscr{B}$ is a Boolean algebra. Let $\mathfrak{A}$ be the σ-algebra generated by the functions f_{B_n}, $B_n \in \mathscr{B}$: it is countably generated by Lemma A.3.

From Step 1 and the Monotone Class Theorem (Proposition 3.1.14, page 85 of [19]), it follows that the family of sets B such that f_B defined in (A.1) is $\mathfrak{A}$-measurable up the an m-negligible set is a σ-algebra containing Ω.

Step 3. Applying the last step to the set $B = h^{-1}(A)$ with $A \in \mathscr{A}$, we obtain that there exists a function f in $L^\infty(m)$, measurable w.r.t. the σ-algebra $\mathfrak{A}$ such that $\chi_A = f$ m-a.e., and this concludes the proof, because up to negligible set f is the characteristic function of a measurable set in $\mathfrak{A}$. □

Remark A.4. We observe there that the result still holds if Ω is the μ-completion of a countably generated σ algebra. More generally, the same proof shows that every σ-algebra $\mathscr{A} \subset \Omega$ is essentially countably generated.

In general, the atoms of $\mathfrak{A}$ are larger than the atoms of $\mathscr{A}$. It is then natural to introduce the following quotient space.

Definition A.5. Let $(\mathtt{A}, \mathscr{A}, m)$ be a measure space, $\mathfrak{A} \subset \mathscr{A}$ a σ-subalgebra. We define the quotient $(\mathtt{L}, \mathscr{L}, \ell)$ as the image space by the equivalence relation

$$\alpha_1 \sim_1 \alpha_2 \quad \Longleftrightarrow \quad \forall A \in \mathfrak{A} \, \big(\alpha_1 \in A \iff \alpha_2 \in A\big).$$

We note that $(\mathscr{L}, \ell)$ is isomorphic as a measure algebra to $(\mathfrak{A}, m)$, so that in the following we will not distinguish the σ-algebras and the measures, but just the spaces $\mathtt{A}$ and $\mathtt{L} = \mathtt{A}/\sim_1$. The quotient map will be denoted by $p : \mathtt{A} \to \mathtt{L}$.

We next define a disintegration of μ *consistent* with the partition $X = \cup_\alpha X_\alpha$ ([10], Definition 452E).

Definition A.6 (Disintegration)**.** The *disintegration* of the probability measure μ *consistent* with the partition $X = \cup_{\alpha \in \mathtt{A}} X_\alpha$ is a map $\mathtt{A} \ni \alpha \mapsto \mu_\alpha \in \mathcal{P}(X, \Omega)$ such that

(1) for all $B \in \Omega$, $\mu_\alpha(B)$ is m-measurable;
(2) for all $B \in \Omega$, $A \in \mathscr{A}$,

$$\mu(B \cap h^{-1}(A)) = \int_A \mu_\alpha(B) m(d\alpha), \tag{A.2}$$

where $h : X \to \mathtt{A}$ is the quotient map and $m = h_\sharp \mu$.

We say that the disintegration is *unique* if for all two measure valued functions $\alpha \mapsto \mu_{1,\alpha}$, $\alpha \mapsto \mu_{2,\alpha}$ which satisfy points (1), (2) it holds $\mu_{1,\alpha} = \mu_{2,\alpha}$ m-a.e. α.

The measures μ_α, $\alpha \in \mathtt{A}$, are called *conditional probabilities*.

We say that the disintegration is *strongly consistent* if for m-a.e. α $\mu_\alpha(X \setminus X_\alpha) = 0$.

We make the following observations.

(1) At this level of generality, we do not require $\mu_\alpha(X_\alpha) = 1$, i.e. that μ_α is concentrated on the class X_α: in fact, we are not even requiring X_α to be μ-measurable.
(2) The choice of the σ-algebra $\mathscr{A}$ in $\mathtt{A}$ is quite arbitrary: in our choice it is the largest σ-algebra which makes point (2) of Definition A.6 meaningful, but one can take smaller σ-algebras, for example Λ considered in Definition A.5.
(3) If $A \in \mathscr{A}$ is an atom of the measure space $(\mathtt{A}, \mathscr{A}, m)$, then the measurability of $\mu_h(B)$ implies that $\mu_h(B)$ is constant m-a.e. on A for all $B \in \Omega$. In particular, if we want to have μ_h concentrated on the smallest possible set, we need to check μ_h with the largest σ-algebra on $\mathtt{A}$: equivalently, this means that the atoms of the measure space $(\mathtt{A}, \mathscr{A}, m)$ are as small as possible. However, negligible sets are useless to this extent.
(4) The formula (A.2) above does not require to have Ω countably generated, and in fact there are disintegration results in general probability spaces (see Section 452 of [10] for general results). However, no general uniqueness result can be expected in that case.
(5) The formula (A.2) can be easily extended to integrable functions by means of monotone convergence theorem: for all μ-integrable functions f, f is μ_α-integrable for m-a.e. α, $\int f \mu_\alpha$ is m-integrable and it holds

$$\int f \mu = \int \bigg(\int f \mu_\alpha \bigg) m(d\alpha). \tag{A.3}$$

We are ready to prove the general disintegration theorem.

Theorem A.7 (Disintegration Theorem)**.** *Assume (X,Ω,μ) countably generated probability space, $X = \cup_{\alpha\in\mathtt{A}} X_\alpha$ a decomposition of X, $h : X \to X_\alpha$ the quotient map. Let $(\mathtt{A},\mathscr{A},m)$ the measure space defined by $\mathscr{A} = h_\sharp \Omega$, $m = h_\sharp \mu$.*

Then there exists a unique disintegration $\alpha \mapsto \mu_\alpha$ consistent with the partition $X = \cup_{\alpha\in\mathtt{A}} X_\alpha$.

Moreover, if $\mathfrak{A}$ is a countably generated σ-algebra such that Proposition A.2 holds, and $\mathtt{L}$ is the quotient space introduced in Definition A.5, $p : \mathtt{A} \to \mathtt{L}$ the quotient map, then the following properties hold:

1. *$X = X_\lambda = (p\circ h)^{-1}(\lambda)$ is μ-measurable, and $X = \cup_{\lambda\in\mathtt{L}} X_\lambda$;*
2. *the disintegration $\mu = \int_{\mathtt{L}} \mu_\lambda m(d\lambda)$ is strongly consistent;*
3. *the disintegration $\mu = \int_{\mathtt{A}} \mu_\alpha m(d\alpha)$ satisfies $\mu_\alpha = \mu_{p(\alpha)}$ m-a.e..*

The last point means that the disintegration $\mu = \int_{\mathtt{A}} \mu_\alpha m(d\alpha)$ has conditional probabilities μ_α constant on each atom of $\mathscr{L}$ in $\mathtt{A}$, precisely given by $\mu_\alpha = \mu_\lambda$ for $\alpha = p^{-1}(\lambda)$ m-a.e.: i.e. μ_α is the pullback of the measure μ_λ.

Proof. We base the proof on well known Disintegration Theorem for measurable functions from $\mathbb{R}^d$ into $\mathbb{R}^{d-k}$, see for example [1], Theorem 2.28.

Step 1: Uniqueness. To prove uniqueness, let $\mathscr{B} = \{B_n\}_{n\in\mathbb{N}}$ be a countable algebra of sets generating Ω. We observe that the $L^\infty(m)$ functions given by $\int_A f_n(\alpha) m(d\alpha) = \mu(h^{-1}(A)\cap B_n)$ are uniquely defined up to a m-negligible set. This means that $\mu_\alpha(B_n)$ is uniquely defined on the algebra $\mathscr{B}$ m-a.e., so that it is uniquely determined on the σ-algebra Ω generated by $\mathscr{B}$.

Step 2: Existence. By measurable space isomorphisms (see for example the proof of the last theorem of [12]), we can consider $(X,\Omega) = (\mathtt{L},\mathfrak{A}) = ([0,1],\mathcal{B})$, so that there exists a unique strongly consistent disintegration $\mu = \int_{\mathtt{L}} \mu_\lambda m(d\lambda)$ by Theorem 2.28 of [1] and Step 1 of the present proof.

Step 3: Point (3) Again by the uniqueness of Step 1, we are left in proving that $\int \mu_{p(\alpha)} m(d\alpha)$ is a disintegration on $X = X_\alpha$.

Since $p : \mathtt{A} \to \mathtt{L}$ is measurable and p is measure preserving, $\alpha \mapsto \mu_{p(\alpha)}(B)$ is m-measurable for all $B \in \Omega$. By Proposition A.2, for all $A \in \mathscr{A}$ there exists $\hat{A} \in \mathscr{L}$ such that $m(A \vartriangle \hat{A}) = \mu(h^{-1}(A) \vartriangle h^{-1}(\hat{A})) = 0$: then

$$\int_{\mathtt{A}} \mu_{p(\alpha)}(B) m(d\alpha) = \int_{\hat{A}} \mu_{p(\alpha)}(B) m(d\alpha) = \int_{\hat{A}} \mu_\lambda(B) m(d\lambda) = \mu(h^{-1}(\hat{A}) \cap B) = \mu(h^{-1}(A)\cap B).$$

□

The final result concerns the existence of a section S for the equivalence relation $X = \cup_\alpha X_\alpha$, under the additional assumption that the atoms of Ω are singletons.

Definition A.8. We say that S is a *section* for the equivalence relation $X = \cup_{\alpha\in\mathtt{A}} X_\alpha$ if for $\alpha \in \mathtt{A}$ there exists a unique $x_\alpha \in S \cap X_\alpha$.

We say that S_μ is a *μ-section* for the equivalence relation induced by the partition $X = \cup_{\alpha\in\mathtt{A}} X_\alpha$ if there exists a Borel set $\Gamma \subset X$ of full μ-measure such that the decomposition

$$\Gamma = \bigcup_{\alpha\in\mathtt{A}} \Gamma_\alpha = \bigcup_{\alpha\in\mathtt{A}} \Gamma \cap X_\alpha$$

has section S_μ.

Clearly from the Axiom of Choice, there is certainly a section S, and by pushing forward the σ-algebra Ω on S we can make $(S,\mathscr{S})$ a measurable space. The following result is a classical application of selection principles.

Proposition A.9. *The disintegration of μ consistent with the partition $X = \cup_{\alpha\in\mathtt{A}} X_\alpha$ is strongly consistent if and only if there exists a Ω-measurable μ-section S such that the σ-algebra $\mathscr{S}$ contains $\mathcal{B}(S)$.*

Proof. Since we are looking for a μ-section, we can replace (X,Ω) with $([0,1],\mathcal{B})$ by a measurable injection.

If the disintegration is strongly consistent, then the map $x \mapsto \{\alpha : x \in X_\alpha\}$ is a μ-measurable map by definition, where the measurable space $(\mathtt{A},\mathscr{A})$ can be taken to be $([0,1],\mathcal{B})$ (Step 2 of Theorem A.7). By removing a set of μ-measure 0, we can assume that h is Borel, so that by Proposition 5.1.9 of [19] it follows that there exists a Borel section.

The converse is a direct consequence of Theorem A.7 and the Isomorphism Theorem among Borel spaces, Theorem 3.3.13 of [19]. □

A.1. **Characterization of the disintegration for a family of equivalence relations.** Consider a family of equivalence relations on X,

$$\mathfrak{E} = \Big\{ E_{\mathsf{e}} \subset X \times X : \ E_{\mathsf{e}} \text{ equivalence relation}, \ \mathsf{e} \in \mathcal{E} \Big\}$$

closed under countable intersection. By Theorem A.7, to each $E_e \in \mathfrak{E}$ we can associate the disintegration

$$X = \bigcup_{\alpha \in \mathtt{A}_e} X_\alpha, \quad \mu = \int_{\mathtt{A}_e} \mu_\alpha m_e(d\alpha), \quad m_e := (h_e)_\sharp \mu, \ h_e : X \to \mathtt{A}_e \text{ quotient map.}$$

The key point of this section is the following easy lemma. For simplicity we will use the language of measure algebras: their elements are the equivalence classes of measurable sets w.r.t. the equivalence relation

$$A \sim A' \quad \Longleftrightarrow \quad \mu(A \vartriangle A') = 0.$$

Let $\mathfrak{Z} = \{\mathscr{C}_{\mathsf{z}}, \mathsf{z} \in \mathcal{Z}\}$ be a family of countably generated σ-algebras such that $\mathscr{C}_{\mathsf{z}} \subset \mathscr{A}$, where $\mathscr{A}$ is a given countably generated σ-algebra on X. Let $\mathscr{C}$ be the σ-algebra generated by $\cup \mathfrak{Z} = \cup_{\mathsf{z} \in \mathcal{Z}} \mathscr{C}_{\mathsf{z}}$.

Lemma A.10. *There is a countable subfamily $\mathfrak{Z}' \subset \mathfrak{Z}$ such that the measure algebra generated by $\mathfrak{Z}'$ coincides with the measure algebra of $\mathscr{C}$.*

Proof. The proof follows immediately by observing that $\mathscr{C}$ is essentially countably generated because it is a σ-subalgebra of $\mathscr{A}$: one can repeat the proof of Proposition A.2, see also Remark A.4, or [9], Proposition 332T(b).

Let A_n, $n \in \mathbb{N}$, be a generating family for $\mathscr{C}$: it follows that there is a countable subfamily $\mathfrak{Z}_n \subset \mathfrak{Z}$ such that A_n belongs to the σ-algebra generated by $\cup \mathfrak{Z}_n = \cup_{\mathscr{C}_{\mathsf{z}} \in \mathfrak{Z}_n} \mathscr{C}_{\mathsf{z}}$. Let $A_{\mathsf{z}m}$, $m \in \mathbb{N}$, be the countable family of sets generating $\mathscr{C}_{\mathsf{z}} \in \mathfrak{Z}$: it is straightforward that $\{A_{\mathsf{z}m}, m \in \mathbb{N}, \mathscr{C}_{\mathsf{z}} \in \cup_n \mathfrak{Z}_n\}$ essentially generates $\mathscr{C}$. □

We can then state the representation theorem.

Theorem A.11. *Assume that the family $\mathfrak{E}$ of equivalence relations is closed w.r.t. countable intersection: if $E_{\mathsf{e}_n} \in \mathfrak{E}$ for all $n \in \mathbb{N}$, then*

$$\bigcap_n E_{\mathsf{e}_n} \in \mathfrak{E}.$$

Then there exists $E_{\bar{\mathsf{e}}} \in \mathfrak{E}$ such that for all E_{e}, $\mathsf{e} \in \mathcal{E}$, the following holds:

1. *if $\mathscr{A}_{\mathsf{e}}$, $\mathscr{A}_{\bar{\mathsf{e}}}$ are the σ-subalgebras of Ω made of the saturated sets for E_{e}, $E_{\bar{\mathsf{e}}}$ respectively, then for all $A \in \mathscr{A}_{\mathsf{e}}$ there is $A' \in \mathscr{A}_{\bar{\mathsf{e}}}$ such that $\mu(A \vartriangle A') = 0$;*
2. *if m_{e}, $m_{\bar{\mathsf{e}}}$ are the restrictions of μ to $\mathscr{A}_{\mathsf{e}}$, $\mathscr{A}_{\bar{\mathsf{e}}}$ respectively, then $\mathscr{A}_{\mathsf{e}}$ can be embedded (as measure algebra) in $\mathscr{A}_{\bar{\mathsf{e}}}$ by point (1): let*

$$m_{\bar{\mathsf{e}}} = \int m_{\bar{\mathsf{e}},\alpha} m_{\mathsf{e}}(d\alpha)$$

be the unique consistent disintegration of $m_{\bar{\mathsf{e}}}$ w.r.t. the equivalence classes of $\mathscr{A}_{\mathsf{e}}$ in $\mathscr{A}_{\bar{\mathsf{e}}}$.

3. *If*

$$\mu = \int \mu_{\mathsf{e},\alpha} m_{\mathsf{e}}(d\alpha), \quad \mu = \int \mu_{\bar{\mathsf{e}},\beta} m_{\bar{\mathsf{e}}}(d\beta)$$

are the unique consistent disintegration w.r.t. E_{e}, $E_{\bar{\mathsf{e}}}$ respectively, then

$$\mu_{\mathsf{e},\alpha} = \int \mu_{\bar{\mathsf{e}},\beta} m_{\bar{\mathsf{e}},\alpha}(d\beta). \tag{A.4}$$

for m_{e}-.a.e. α.

The last point essentially tells us that the disintegration w.r.t. $E_{\bar{\mathsf{e}}}$ is the sharpest one, the others being obtained by integrating the conditional probabilities $\mu_{\bar{\mathsf{e}},\beta}$ w.r.t. the probability measures $m_{\bar{\mathsf{e}},\alpha}$.

Note that the result is useful but it can lead to trivial result if $E = \{(x,x), x \in X\}$ belongs to $\mathfrak{E}$: in this case

$$\mu_{\bar{\mathsf{e}},\beta} = \delta_\beta, \quad m_{\bar{\mathsf{e}},\alpha} = \mu_{\mathsf{e},\alpha}.$$

Proof. Point (1). We first notice that if $E_{\mathsf{e}_1}, E_{\mathsf{e}_2} \in \mathfrak{E}$ and $A \in \mathscr{A}_{\mathsf{e}_1}$, the σ-algebra of saturated sets generated by E_{e_1}, then $A \in \mathscr{A}_{\mathsf{e}_{12}}$, the σ-algebra of saturated sets generated by $E_{\mathsf{e}_1} \cap E_{\mathsf{e}_2}$. Hence, if $\mathfrak{E}$ is closed under countable intersection, then for every family of equivalence relations $E_{\mathsf{e}_n} \in \mathfrak{E}$ there exists $E_{\bar{\mathsf{e}}} \in \mathfrak{E}$ such that the σ-algebras $\mathscr{A}_{\mathsf{e}_n}$ made of the saturated measurable sets w.r.t. E_{e_n} are σ-subalgebras of the σ-algebra $\mathscr{A}_{\bar{\mathsf{e}}}$ made of the saturated measurable sets w.r.t. the equivalence relation $E_{\bar{\mathsf{e}}}$.

By Lemma A.10 applied to the family $\mathfrak{Z} = \{\mathscr{A}_{\mathsf{e}} | \mathsf{e} \in \mathcal{E}\}$, we can take a countable family of equivalence relations such that the σ-algebra of saturated sets w.r.t. their intersection satisfies Point (1).

Point (2). This point is a consequence of the Disintegration Theorem A.7, using the embedding $\mathscr{A}_{\mathsf{e}} \ni A \mapsto A' \in \mathscr{A}_{\bar{\mathsf{e}}}$ given by the condition $\mu(A \vartriangle A') = 0$.

Point (3). Since consistent disintegrations are unique, it is enough to show that (A.4) is a disintegration for E_{e}. By definition, for all $C \in \Omega$, $\mu_{\bar{\mathsf{e}},\beta}(C)$ is a $m_{\bar{\mathsf{e}}}$-measurable function, so that by (A.3) it is also $m_{\bar{\mathsf{e}},\alpha}$-measurable for m_e-a.e. α and

$$\alpha \mapsto \int \mu_{\bar{\mathsf{e}},\beta}(C) m_{\bar{\mathsf{e}},\alpha}(d\beta)$$

is m_{e}-measurable. Denoting with h_{e} the equivalence map for E_{e}, for all $A \in \mathscr{A}_{\mathsf{e}}$ we have

$$\mu(C \cap h_{\mathsf{e}}^{-1}(A)) = \int_A \mu_{\bar{\mathsf{e}},\beta}(C) m_{\bar{\mathsf{e}}}(d\beta) = \int_A \bigg(\int \mu_{\bar{\mathsf{e}},\beta}(C) m_{\bar{\mathsf{e}},\alpha}(d\beta) \bigg) m_{\mathsf{e}}(d\alpha),$$

where we used the definition of $\mu_{\bar{\mathsf{e}},\beta}$ in the first equality and (A.3) in the second one. □

Remark A.12. Using the fact that if $A \in \mathscr{A}_e$, then there exists $A' \in \mathscr{A}_{\bar{e}}$ such that $\mu(A \vartriangle A') = 0$, then the identity map $\mathbb{I} : (X, \mathscr{A}_{\bar{e}}) \to (X, \mathscr{A}_e)$ is μ-measurable. Let $h_e : X \to X/E$, $h_{\bar{e}} : X \to X/\bar{E}$ be quotient maps. It is fairly easy to show that there exists a unique $(h_{\bar{e}})_\sharp \mu$-measurable map $g_{\bar{e}e}$ (defined up to negligible sets) such that the following diagram commutes:

$$\begin{tikzcd} X \arrow[r, "h_{\bar{e}}"] \arrow[d, "\mathbb{I}"'] & X/E_{\bar{e}} \arrow[d, "g_{\bar{e}e}"] \\ X \arrow[r, "h_e"'] & X/E_e \end{tikzcd} .$$

Clearly one then has

$$(h_e)_\sharp \mu = (g_{\bar{e}e})_\sharp \big((h_{\bar{e}})_\sharp \mu \big) = (g_{\bar{e}e} \circ h_{\bar{e}})_\sharp \mu. \tag{A.5}$$

Appendix B. Perturbation by cycles

For the particular applications we are considering, the geometrical constraints allow only to perturb a given measure $\pi \in \Pi(\mu, \nu)$ by means of bounded measures $\lambda \in \mathcal{M}([0,1]^2)$ with 0 marginals, and such that $\pi + \lambda \geq 0$. The simplest way of doing this perturbation is to consider closed cycles in $[0,1]^2$: we will call this types of perturbation *perturbation by cycles* (a more precise definition is given below).

The problem of checking whether a measure μ can be perturbed by cycles has been considered in several different contexts, see for example [2, 3, 11]. Here we would like to construct effectively a perturbation, which will be (by definition) a perturbation by cycles.

Since we are using a duality result valid only for analytic sets, in the following we will restrict to a coanalytic cost c. We first recall some useful results on analytic subsets of Polish spaces (in our case $[0,1]$), and the main results of [13].

B.1. **Borel, analytic and universally measurable sets.** Our main reference is [19].

The *projective class* $\Sigma^1_1(X)$ is the family of subsets A of the Polish space X for which there exists Y Polish and $B \in \mathcal{B}(X \times Y)$ such that $A = P_1(B)$. The *coprojective class* $\Pi^1_1(X)$ is the complement in X of the class $\Sigma^1_1(X)$. The σ-algebra generated by Σ^1_1 is denoted by $\mathcal{A}$.

The *projective class* $\Sigma^1_{n+1}(X)$ is the family of subsets A of the Polish space X for which there exists Y Polish and $B \in \Pi^1_n(X \times Y)$ such that $A = P_1(B)$. The *coprojective class* $\Pi^1_{n+1}(X)$ is the complement in X of the class $\Sigma^1_{n+1}(X)$.

If Σ^1_n, Π^1_n are the projective, coprojective pointclasses, then the following holds (Chapter 4 of [19]):

(1) Σ^1_n, Π^1_n are closed under countable unions and countable intersections (in particular they are monotone classes);

(2) Σ^1_n is closed w.r.t. projections, Π^1_n is closed w.r.t. coprojections;
(3) the *ambiguous class* $\Delta^1_n := \Sigma^1_n \cap \Pi^1_n$ is a σ-algebra and $\Sigma^1_n \cup \Pi^1_n \subset \Delta^1_{n+1}$.

We recall that a subset of X Polish is *universally measurable* if it belongs to all completed σ-algebras of all Borel measures on X: it can be proved that every set in $\mathcal{A}$ is universally measurable.

Under the axiom of Projective Determinacy (PD) all projective sets are universally measurable, and PD is undecidable in ZFC ([15, 16]). *In the rest of the present Appendix we assume* (PD). One could avoid this assumption by recovering independently the measurability of the functions we are going to define by countable limit procedures (see for example [6]), but since our aim is to describe a construction this analysis is not needed here.

We recall that Borel counterimages of universally measurable sets are universally measurable.

B.2. **General duality results.** All the results recalled in this sections are contained in [13].

Let $A \subset [0,1]^d$ be a subset of $[0,1]^d$, and consider Borel probabilities $\mu_i \in \mathcal{P}([0,1])$, $i = 1,\dots,d$. We want to know if there is a measure π such that $\pi^*(A) > 0$ and its marginals are bounded by the measure μ_i: $(P_i)_\sharp \pi \leq \mu_i$. We recall that

$$\pi^*(A) := \inf\big\{\pi(A') : \ A' \in \mathcal{B}([0,1]^d), A \subset A'\big\} \tag{B.1}$$

is the outer π measure. For simplicity, we will denote the i-th measure space in the product with X_i.

Definition B.1. A set $A \subset [0,1]^d$ is *cross-negligible w.r.t. the measures* μ_i, $i = 1,\dots,d$, if there are μ_i-negligible sets N_i, $i = 1,\dots,d$, such that $A \subset \cup_i P_i^{-1}(N_i)$.

Given $A_1, A_2 \in [0,1]^d$, we define

$$\mathrm{dist}(A_1,A_2) := \inf\bigg\{\sum_{i=1}^d \int h_i\mu_i : \ \chi_{A_1 \vartriangle A_2}(x) \leq \sum_{i=1}^d h_i(x_i), \ h_i \in L^\infty(\mu_i)\bigg\}. \tag{B.2}$$

We say that $A_1, A_2 \subset [0,1]^d$ are *equivalent* and we write $A_1 \sim_{\mathrm{dist}} A_2$ if $A_1 \vartriangle A_2$ is cross negligible, i.e. $\mathrm{dist}(A_1,A_2) = 0$.

This definition is the same as the L-shaped sets defined in [3]. Clearly the cross-negligible sets can be taken to be G_δ-sets. The fact that $\sim_{\mathrm{dist}}$ is an equivalence relation and that $(\mathbf{P}([0,1]^d)/\sim_{\mathrm{dist}}, \mathrm{dist})$ is a metric space is proved in [13], Proposition 1.15. Following again [13], given $\mathcal{A} \subset \mathbf{P}([0,1]^d)$, we denote $\overline{\mathcal{A}}$ the closure of $\mathcal{A}$ w.r.t. the distance dist.

The next theorem collects some of the main results of [13]. This results are duality results, which compare the supremum of a linear function in the convex set

$$\Pi(\mu_1,\dots,\mu_d) := \Big\{\pi \in \mathcal{P}([0,1]^d) : \ (P_i)_\sharp \pi = \mu_i, i = 1,\dots,d\Big\}$$

with the infimum of a convex function in a predual space.

Theorem B.2. *If $A \in \overline{\Sigma^1_1(\mathbb{R}^d)}$, then the following duality holds*

$$\sup\Big\{\pi(A) : \ \pi \in \Pi(\mu_1,\dots,\mu_d)\Big\} = \min\bigg\{\sum_{i=1}^d \int h_i\mu_i : \ \sum_{i=1}^d h_i(x_i) \geq \chi_A(x), 0 \leq h_i \leq 1\bigg\}. \tag{B.3}$$

Moreover, if A is in closure w.r.t. d of the family of closed sets, then the max on the l.h.s. is reached. In particular the maximum is reached when A is in the class of countable intersections of elements of the product algebra.

Proof. The fact that the duality (B.3) holds with the infimum in the r.h.s. is a consequence of [13], Theorem 2.14. In our settings the analytic sets contain all the Borel sets, so that in particular the duality holds for Borel sets.

The fact that the minimum is reached is a consequence of [13], Theorem 2.21.

Finally, the last assertion follows from [13], Theorem 2.19, and the subsequent remarks. □

A fairly easy corollary is that if the supremum of (B.3) is equal to 0, then A is cross negligible.

Remark B.3. Note that since we are considering a maximum problem for a positive linear functional, then the problem is equivalent when considered in the larger space

$$\Pi^{\leq}(\mu_1,\dots,\mu_d) := \Big\{0 \leq \pi \in \mathcal{M}([0,1]^d) : \ (P_i)_\sharp \pi \leq \mu_i, i = 1,\dots,d\Big\}.$$

B.3. **Decomposition of measures with 0 marginals.** In this section we decompose a measure with 0 marginals into its cyclic or essentially cyclic part and acyclic part. The decomposition is not unique, even if we can determine if a perturbation is cyclic, essentially cyclic or acyclic.

Let Λ be the convex closed set of Borel measures on $[0,1]^2$ with 0 marginals:

$$\Lambda := \Big\{ \lambda \in \mathcal{M}([0,1]^d) : \ (P_i)_\sharp \lambda = 0, \ i = 1,2 \Big\}. \tag{B.4}$$

In the following we restrict to $d = 2$, in view of applications to the transport problem in $[0,1]^2$.

Definition B.4. We define the following sets.

The *configuration set*

$$C_n := \Big\{ w \in [0,1]^{2n} : \ P_{2i-1} w \neq (P_{2i+1 \mod 2n}) w, \ P_{2i} w \neq (P_{2i+2 \mod 2n}) w, i = 1, \dots, n \Big\}.$$

The *phase set*

$$D_n := \Big\{ z \in [0,1]^{4n} : \ (P_{4i-1}, P_{4i}) z = (P_{4i+1 \mod 4n}, P_{4i-2}) z, i = 1, \dots, n \Big\}.$$

The *set of finite cycles of arbitrary length D_∞*

$$D_\infty := \Big\{ z \in [0,1]^{2\mathbb{N}} : (P_{4i-1}, P_{4i}) z = (P_{4i+1}, P_{4i-2}) z, \exists k : P_{4k+i} z = P_i z, i \in \mathbb{N} \Big\}.$$

The *projection operator*

$$q : [0,1]^{4n} \to [0,1]^{2n}, \quad (P_{2i-1}, P_{2i}) q(z) = (P_{4i-3}, P_{4i-2}) z, \ i = 1, \dots, n.$$

The *reduced phase set*

$$\tilde{D}_n := q^{-1}(C_n) \cap D_n.$$

The *narrow configuration set* and *narrow phase space*

$$\hat{C}_n := \Big\{ w \in [0,1]^{2n} : \ (P_{2i-1}, P_{2i}) w \neq (P_{2j-1}, P_{2k}) w, i \neq j, k \Big\}, \quad \hat{D}_n := q^{-1}(\hat{C}_n) \cap D_n. \tag{B.5}$$

Remark B.5. The following remarks are straightforward.

1. The set C_n is open not connected in $[0,1]^{2n}$, and its connected components are given by the family of sets

   $$C_{n,I} := \Big\{ w \in [0,1]^{2n} : \ P_{2i-1} w \gtrless (P_{2i+1 \mod 2n}) w, \ P_{2i} w \gtrless (P_{2i+2 \mod 2n}) w, i = 1, \dots, n \Big\}$$

   for the 4 possible choices of the inequalities and of $i \in \{1, \dots, n\}$.
2. The set D_n is compact connected, and the set $\tilde{D}_n$ can be written as

   $$\begin{aligned} \tilde{D}_n := \Big\{ z \in [0,1]^{4n} : \ & (P_{4i-1}, P_{4i}) z = (P_{4i+1 \mod 4n}, P_{4i-2 \mod 4n}) z, \\ & P_{4i-3} z \neq (P_{4i+1 \mod 4n}) z, \ P_{4i-2} z \neq (P_{4i+2 \mod 4n}) z, i = 1, \dots, n \Big\}. \end{aligned}$$
3. Both sets C_n and D_n are invariant for the cyclical permutation of coordinates T defined by $(P_{i+2 \mod n})(Tw) = P_i w$, $i = 1, \dots, 2n$ in $[0,1]^{2n}$ and by $q^{-1} T q$ on D_n.
4. The narrow phase set is made by cycles of length exactly n.

We give now the following definitions.

Definition B.6. A measure λ is *n-cyclic* if there exists $m \in \mathcal{M}^+(C_n)$ such that

$$\lambda^+ = \frac{1}{n} \int_{C_n} \sum_{i=1}^n \delta_{P_{(2i-1,2i)} w} m(dw), \quad \lambda^- = \frac{1}{n} \int_{C_n} \sum_{i=1}^n \delta_{P_{(2i+1,2i \mod 2n)} w} m(dw). \tag{B.6}$$

A n-cyclic measure λ is a *simple n-cycle* if m is supported on a set $q(Q)$ with

$$Q = \Big\{ z \in D_n : (P_{4i-3}, P_{4i-2}) z \in (x_i, y_i) + [-\epsilon, \epsilon]^2, \min_{i,j} \big\{ |x_i - x_j|, |y_i - y_j| \big\} \geq 2\epsilon \Big\}.$$

A measure λ is *cyclic* if there exist $m_n \in \mathcal{M}^+(C_n)$, $n \in \mathbb{N}$, such that $\sum_n m_n(C_n) < \infty$ and

$$\lambda^+ = \sum_n \frac{1}{n} \int_{C_n} \sum_{i=1}^n \delta_{P_{(2i-1,2i)} w} m_n(dw), \quad \lambda^- = \sum_n \frac{1}{n} \int_{C_n} \sum_{i=1}^n \delta_{P_{(2i+1 \mod 2n, 2i)} w} m_n(dw). \tag{B.7}$$

From the definition of simple n-cycles it follows that there are disjoint $2n$ sets $(x_i, y_i) + [-\epsilon, \epsilon]^2$, $(x_{i+1 \mod n}, y_i) + [-\epsilon, \epsilon]^2$, $i = 1, \dots, n$, such that

$$\lambda^+ \bigg(\bigcup_{i=1}^n (x_i, y_i) + [-\epsilon, \epsilon]^2 \bigg) + \lambda^- \bigg(\bigcup_{i=1}^n (x_{i+1 \mod n}, y_i) + [-\epsilon, \epsilon]^2 \bigg) = \|\lambda\|.$$

The next lemma is a simple consequence of the separability of $[0,1]^{4n}$ and the fact that $\hat{C}_n$ is open.

Lemma B.7. *Each n-cyclic measure λ of the form*

$$\lambda^+ = \frac{1}{n} \int_{\hat{C}_n} \sum_{i=1}^n \delta_{P_{(2i-1,2i)} w} m(dw), \quad \lambda^- = \frac{1}{n} \int_{\hat{C}_n} \sum_{i=1}^n \delta_{P_{(2i+1,2i \mod 2n)} w} m(dw)$$

can be written as the sum of simple n-cycles λ_i so that

$$\lambda^+ = \sum_i \lambda_i^+, \quad \lambda^- = \sum_i \lambda_i^-.$$

B.3.1. *n-cyclic components of a measure.* Consider the Jordan decomposition of $\lambda \in \Lambda$,

$$\lambda = \lambda^+ - \lambda^- \qquad \lambda^+ \perp \lambda^-, \ \lambda^+, \lambda^- \geq 0,$$

and the Borel sets A^+, A^- of the Hahn decomposition:

$$A^+ \cap A^- = \emptyset, \ A^+ \cup A^- = [0,1]^2, \qquad \lambda^+ = \lambda \llcorner_{A^+}, \ \lambda^- = \lambda \llcorner_{A^-}.$$

Define then

$$\mu_{2i-1} := \lambda^+, \quad \mu_{2i} := \lambda^-, \tag{B.8}$$

with $i = 1, \dots, n$.

From Theorem B.2 and the fact that D_n is compact, the following proposition follows.

Proposition B.8. *Let μ_i as in* (B.8)*. There exists a solution to the marginal problem, for $n \in \mathbb{N}$,*

$$\max \Big\{ \pi(\tilde{D}_n) : \ \pi \in \Pi(\mu_1, \dots, \mu_{2n}) \Big\} = \min \bigg\{ \sum_{i=1}^{2n} \int_{[0,1]^2} h_i \mu_i : \ \sum_{i=1}^{2n} h_i((P_{2i-1}, P_{2i})z) \geq \chi_{\tilde{D}_n}(z) \bigg\}. \tag{B.9}$$

Proof. It is enough to prove that D_n is in the equivalence class of $\tilde{D}_n$ w.r.t. $\sim_{\mathrm{dist}}$: from this it follows that for every measure in $\Pi(\mu_i)$ one has $\pi(D_n) = \pi(\tilde{D}_n)$, and then one can apply Theorem B.2.

Step 1. By definition

$$D_n \setminus \tilde{D}_n \subset \bigcup_{i=1}^n \Big\{ z : P_{4i-3} z = (P_{4i+1 \mod 4n}) \text{ or } P_{4i-2} z = (P_{4i+2 \mod 4n}) z \Big\},$$

so that if $z \in D_n \setminus \tilde{D}_n$ for at least one i

$$(P_{4i-3}, P_{4i-2})z = (P_{4i-1}, P_{4i})z \quad \text{or} \quad (P_{4i-1}, P_{4i})z = (P_{4i+1 \mod 4n}, P_{4i+2 \mod 4n})z. \tag{B.10}$$

Step 2. Consider the functions, for $i = 1, \dots, n$,

$$f_{2i-1} = \chi_{[0,1]^2 \setminus A^+}, \quad f_{2i} = \chi_{[0,1]^2 \setminus A^-}.$$

Since $f_{2i-1} + f_{2i} \geq 1$, it follows from (B.10) that

$$\sum_{i=1}^n f_{2i-1}\big((P_{4i-3}, P_{4i-2})z\big) + f_{2i}\big((P_{4i-1}, P_{4i})z\big) \geq \chi_{D_n \setminus \tilde{D}_n}.$$

Step 3. Since $\lambda^+(A^-) = \lambda^-(A^+) = 0$, then

$$\sum_{i=1}^n \int_{[0,1]^2} f_{2i-1} \mu_{2i-1} + \int_{[0,1]^2} f_{2i} \mu_{2i} = \sum_{i=1}^n \lambda^+(A^-) + \lambda^-(A^+) = 0.$$

Hence $\mathrm{dist}(D_n, \tilde{D}_n) = 0$. □

We now define the n-cyclic components of λ.

Definition B.9. Let π be a maximizer for (B.9) and define the measure

$$\lambda_n := \frac{1}{n} \sum_{i=1}^{n} (P_{4i-3}, P_{4i-2})_\sharp \pi \llcorner_{\tilde{D}_n} - \frac{1}{n} \sum_{i=1}^{n} (P_{4i-1}, P_{4i})_\sharp \pi \llcorner_{\tilde{D}_n}.$$

We say that λ_n is the (or better a) *n-cyclic component* of λ.

Remark B.10. The following are easy remarks.

(1) $0 \leq \lambda_n^+ \leq \lambda^+$ and $0 \leq \lambda_n^- \leq \lambda^-$: in fact, by construction

$$0 \leq (P_{4i-3}, P_{4i-2})_\sharp \pi \llcorner_{\tilde{D}_n} \leq \lambda^+, \quad 0 \leq (P_{4i-1}, P_{4i})_\sharp \pi \llcorner_{\tilde{D}_n} \leq \lambda^+.$$

Moreover, by the definition of D_n, it follows that

$$\left| (P_{4i-3}, P_{4i-2})_\sharp \pi \llcorner_{\tilde{D}_n} \right| = \left| (P_{4i-1}, P_{4i})_\sharp \pi \llcorner_{\tilde{D}_n} \right| = \pi(D_n),$$

so that $|\lambda_n| = 2\pi(D_n)$.

(2) If π is a maximum, also the symmetrized measure

$$\tilde{\pi} := \frac{1}{n} \sum_{i=0}^{n-1} \Big(\underbrace{T \circ \cdots \circ T}_{i-\text{times}} \Big)_\sharp \pi$$

is still a maximum. For this measure $\tilde{\pi}$ it follows that

$$\lambda_n = (P_{4i-3}, P_{4i-2})_\sharp \tilde{\pi} \llcorner_{\tilde{D}_n} - (P_{4i-1}, P_{4i})_\sharp \tilde{\pi} \llcorner_{\tilde{D}_n} \tag{B.11}$$

for all $i = 1, \dots, n$. In particular, if we consider again the problem (B.9) with $\lambda_n^\pm$ as marginals in (B.8), then $\tilde{\pi}$ is still a maximum. However, there are maxima which are not symmetric, and for which the projection on a single component does not exhibit a cyclic structure, as in Example B.17.

(3) The n-cyclic component of λ is a n-cyclic measure, as one can see by the trivial disintegration

$$\pi = \int_{C_n} \delta_{q^{-1}(w)} m(w), \qquad m(w) := (q_\sharp \pi)(w).$$

Conversely, if λ is n-cyclic, then $\pi \llcorner_{D_n} = (q^{-1})_\sharp m$ is a maximum for the problem (B.9).

Note that the condition

$$\lambda = \frac{1}{n} \int_{C_n} \sum_{i=1}^{n} \Big(\delta_{P_{(2i-1,2i)} w} - \delta_{P_{(2i+1 \mod 2n, 2i)} w} \Big) m(dw)$$

is not sufficient, because of cancellation, as it can be easily seen by the measure

$$\lambda = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}.$$

(4) If $\lambda_n = 0$, it follows from the duality stated in Theorem B.2 that D_n is cross negligible, so that there exists Borel sets N_i, $i = 1, \dots, n$ such that

$$\lambda^+(N_{2i-1}) = \lambda^-(N_{2i}) = 0 \quad \text{and} \quad D_n \subset \bigcup_{i=1}^{2n} (P_i)^{-1}(N_i).$$

Hence the sets

$$N^+ = \bigcup_{i=1}^{n} N_{2i-1}, \quad N^- = \bigcup_{i=1}^{n} N_{2i}$$

still satisfy $\lambda^+(N^+) = \lambda^-(N^-) = 0$ and

$$D_n \cap \bigcap_{i=1}^{n} (P_{2i-1})^{-1}(N^+)^c \cap (P_{2i})^{-1}(N^-)^c = \emptyset.$$

We thus conclude that if $\lambda_n = 0$ there exist Borel sets A^+, A^- such that λ^+ is concentrated in A^+, λ^- is concentrated in A^- and there is no n-cycle $\{(x_i, y_i), i = 1, \dots, n\}$ such that $(x_i, y_i) \in A^+$ and $(x_{i+1 \mod n}, y_i) \in A^-$ for all $i = 1, \dots, n$.

Define the measure $\lambda_{\not{n}} := \lambda - \lambda_n$.

Lemma B.11. *The n-cyclic component of $\lambda_{\not{n}}$ is zero. Equivalently, $\lambda_{\not{n}}$ satisfies*

$$\max\Big\{\pi(\tilde{D}_n), \pi \in \Pi(\mu_1, \dots, \mu_{2n})\Big\} = 0 \tag{B.12}$$

for the marginal problem

$$\mu_i = \begin{cases} \lambda^+_{\not{n}} & i \text{ odd} \\ \lambda^-_{\not{n}} & i \text{ even} \end{cases}.$$

Proof. If in (B.12) we have a positive maximum π', then we can assume this maximum to be symmetric, so that (B.11) holds. Let π be a symmetric positive maximum of the original (B.9): by construction we have that

$$0 \le \lambda_n^+ = (P_{(1,2)})_\sharp \pi \llcorner_{D_n} \le \lambda^+, \quad 0 \le \lambda_n^- = (P_{(3,4)})_\sharp \pi \llcorner_{D_n} \le \lambda^-$$
$$0 \le (P_{(1,2)})_\sharp \pi' \le \lambda^+ - \lambda_n^+, \quad 0 \le (P_{(3,4)})_\sharp \pi' \le \lambda^- - \lambda_n^-,$$

so that

$$0 \le \lambda_n^+ + (P_{(1,2)})_\sharp \pi' = (P_{(1,2)})_\sharp (\pi + \pi') \le \lambda^+, \quad 0 \le \lambda_n^- + (P_{(3,4)})_\sharp \pi' = (P_{(3,4)})_\sharp (\pi + \pi') \le \lambda^-,$$

and $(\pi + \pi')(D_n) > \pi(D_n)$, contradicting the maximality of π. □

A measure can be decomposed into a *cyclic and an acyclic part* by removing n-cyclic components for all $n \in \mathbb{N}$ (see Remark B.13). However, when removing a n-cyclic component the m-cyclic components are affected, for $m \neq n$. More clearly, the following observations are in order.

For all $n, k \in \mathbb{N}$ one has

$$\max\Big\{\pi(\tilde{D}_n), \pi \in \Pi(\mu_1, \dots, \mu_{2n})\Big\} \le \max\Big\{\pi(\tilde{D}_{kn}), \pi \in \Pi(\mu_1, \dots, \mu_{2kn})\Big\},$$

because if π_1 is a measure in $\Pi(\mu_1, \dots, \mu_{2n})$, then the measure

$$\pi_2 = \big(\underbrace{\mathbb{I}_n, \dots, \mathbb{I}_n}_{k-\text{times}}\big)_\sharp \pi_1$$

belongs to $\Pi(\mu_1, \dots, \mu_{2kn})$ and $\pi_2(\tilde{D}_{kn}) = \pi_2(D_{kn}) = \pi_1(D_n) = \pi_1(\tilde{D}_n)$.

However, in general

$$\max\Big\{\pi(\tilde{D}_n), \pi \in \Pi(\mu_1, \dots, \mu_{2n})\Big\} + \max\Big\{\pi(\tilde{D}_n), \pi \in \Pi(\nu_1, \dots, \nu_{2kn})\Big\}$$
$$< \max\Big\{\pi(\tilde{D}_{kn}), \pi \in \Pi(\mu_1, \dots, \mu_{2kn})\Big\},$$

where we define

$$\nu_i = \begin{cases} \lambda^+_{\not{n}} & i \text{ odd} \\ \lambda^-_{\not{n}} & i \text{ even} \end{cases} \qquad \mu_i = \begin{cases} \lambda^+ & i \text{ odd} \\ \lambda^- & i \text{ even} \end{cases}.$$

This can be seen in Example B.17, by taking $n = 2$ and $k = 4$: in fact for any choice of the maximal solution for $n = 2$ the remaining measure $\lambda - \lambda_{\not{2}}$ does not contain any cycle of length 8, while λ itself is a cycle of length 8. It follows

$$\max\Big\{\pi(\tilde{D}_n): \ \pi \in \Pi(\mu_1, \dots, \mu_{2n})\Big\} + \max\Big\{\pi(\tilde{D}_n): \ \pi \in \Pi(\nu_1, \dots, \nu_{2kn})\Big\} = 2$$
$$< 8 = \max\Big\{\pi(\tilde{D}_{kn}): \ \pi \in \Pi(\mu_1, \dots, \mu_{2kn})\Big\}.$$

An even more interesting example is provided in Example B.18, where it is shown that a measure can be decomposed into a cyclic and an acyclic part in different ways, and the mass of each part depends on the decomposition one chooses.

B.3.2. *Cyclic and essentially cyclic measures.* Given a sequence of marginals μ_i, let

$$\Pi_\infty(\{\mu_i\}_i) = \Big\{\pi \in \mathcal{P}([0,1]^{2\mathbb{N}}) : (P_i)_\sharp \pi = \mu_i, i \in \mathbb{N}\Big\},$$

Consider following problem in $[0,1]^{2\mathbb{N}}$:

$$\sup \Big\{\pi(D_\infty), (P_{2i-1})_\sharp \pi = \lambda^+, (P_{2i})_\sharp \pi = \lambda^-, i \in \mathbb{N}\Big\}.$$

Definition B.12. We say that a measure $\lambda \in \Lambda$ is *essentially cyclic* if

$$\sup \Big\{\pi(D_\infty), (P_{2i-1})_\sharp \pi = \lambda^+, (P_{2i})_\sharp \pi = \lambda^-, i \in \mathbb{N}\Big\} = \lambda^+([0,1]^2) = \lambda^-([0,1]^2).$$

It is clear that if λ is cyclic, then the maximum exists, and viceversa (Remark B.10, Point (3), observing that $D_n \hookrightarrow D_\infty$). If λ is acyclic, then the supremum is equal to 0. Since D_∞ is not closed in $[0,1]^{2\mathbb{N}}$, we cannot state that such a maximum exists.

Remark B.13. We now construct a special decomposition, whose cyclic part however is not necessary maximal.

Define recursively the marginal problem in D_n by

$$\mu_{2n-1} := \lambda^+ - \sum_{i=2}^{n-1} \lambda_i^+ \quad \mu_{2n} := \lambda^- - \sum_{i=2}^{n-1} \lambda_i^-, \tag{B.13}$$

where λ_i is given at the i-th step and λ_n is obtained by

$$\lambda_n := \frac{1}{n} \sum_{i=1}^{n} (P_{4i-3}, P_{4i-2})_\sharp \pi \llcorner_{\tilde{D}_n} - \frac{1}{n} \sum_{i=1}^{n} (P_{4i-1}, P_{4i})_\sharp \pi \llcorner_{\tilde{D}_n}$$

solving the problem

$$\max \Big\{\pi(\hat{D}_n), \pi \in \Pi(\mu_1, \dots, \mu_{2n})\Big\} = \min \bigg\{ \sum_{i=1}^{2n} \int_{[0,1]^2} h_i(x)\mu_i,\ \sum_{i=1}^{2n} h_i((P_{2i-1}, P_{2i})z) \geq \chi_{\hat{D}_n} \bigg\}. \tag{B.14}$$

Let $\{\pi_n\}_{n\in\mathbb{N}}$ be the sequence of maxima for (B.14). There is a canonical way to embed π_n in $\Pi_\infty(\{\mu_i\}_i)$, with μ_i given by (B.13) for $i \in \mathbb{N}$ (here we assume that $\|\lambda\| = 2$). In fact, it is enough to take

$$T_n : [0,1]^{4n} \to [0,1]^{2\mathbb{N}}, \quad z \mapsto T_n(z) = (z, z, z, \dots), \quad \tilde{\pi}_n = (T_n)_\sharp \pi_n.$$

Hence the measure $\tilde{\pi} = \sum_n \tilde{\pi}_n$ belongs to π_∞, the series being strongly converging, and since every map T_n takes values in D_∞, the measure $\tilde{\pi}$ satisfies

$$\tilde{\pi}(D_\infty) = \sum_n \pi_n(D_n).$$

Example B.18 implies that in general $\tilde{\pi}$ it is not a supremum.

Similarly, the measures $\sum_i^n \lambda_i^+$, $\sum_i^n \lambda_i^-$ are strongly convergent to measures λ_c^+, λ_c^-.

The sets D_n are cross negligible for the marginals

$$\mu_i = \begin{cases} \lambda_a^+ = \lambda^+ - \lambda_c^+ & i \text{ odd} \\ \lambda_a^- = \lambda^- - \lambda_c^- & i \text{ even} \end{cases}$$

This follows easily from (B.12) and the fact that the series of λ_n is converging.

Hence, from Point (4) of Remark B.10, one concludes that λ_a^+, λ_a^- are supported on two disjoint sets A_a^+, A_a^-, respectively, so that there are no closed cycles $\{(x_i, y_i), i = 1, \dots, n\}$, $n \in \mathbb{N}$, such that $(x_i, y_i) \in A_a^+$ and $(x_{i+1}, y_i) \in A_a^-$ for all $i = 1, \dots, n$ and $(x_{n+1}, y_{n+1}) = (x_1, y_1)$.

B.3.3. *Perturbation of measures.* For a measure $\pi \in \mathcal{P}([0,1]^2)$ and an analytic set $A \subset [0,1]^2$ such that $\pi(A) = 1$, we give the following definition.

Definition B.14. A *cyclic perturbation on* A of the measure π is a cyclic, nonzero measure λ concentrated on A and such that $\lambda^- \leq \pi$.

Proposition B.15. *If there is no cyclic perturbation of* π *on* A*, then there is* Γ *with* $\pi(\Gamma) = 1$ *such that for all finite sequences* $(x_i, y_i) \in \Gamma$*,* $i = 1, \dots, n$*, with* $x_i \neq x_{i+1 \mod n}$ *and* $y_i \neq y_{i+1 \mod n}$ *it holds*

$$\Big\{(x_{i+1}, y_i), i = 1, \dots, n, x_{n+1} = x_1\Big\} \not\subset A.$$

Proof. Define the set on n-cycles in A as

$$C_{n,A} := q\bigg(D_n \cap \prod^{2n} A\bigg).$$

The fact that there is no cyclic perturbation means that for all $n \in \mathbb{N}$

$$\sup \Big\{m(q(D_n \cap \Pi^{2n} A)) : m \in \Pi(\pi, \dots, \pi)\Big\} = 0.$$

Then $C_{n,A}$ is cross-negligible by Theorem B.2: there exist π-negligible sets N_i such that

$$\bigg(\prod^n \Gamma \setminus (\cup_i N_i)\bigg) \cap C_{n,A} = \emptyset.$$

The set $\Gamma \setminus (\cup_i N_i)$ satisfies the statement of the proposition. □

Let $c \geq 0$ be a Π^1_1-cost function: in the next proposition we follow the ideas of [3], which reduce to ones in [17] for atomic marginals (see [13] for the general result).

Proposition B.16. *If there is no cyclic perturbation* λ *of* π *such that* $\mathcal{I}(\pi + \lambda) < \mathcal{I}(\pi)$*, then there is* Γ *with* $\pi(\Gamma) = 1$ *such that for all finite sequences* $(x_i, y_i) \in \Gamma$*,* $i = 1, \dots, I$*,* $x_{I+1} := x_1$ *it holds*

$$\sum_{i=1}^{I} \big[c(x_{i+1}, y_i) - c(x_i, y_i)\big] \geq 0. \tag{B.15}$$

Proof. Let Γ a σ-compact carriage of π such that $c \llcorner \Gamma$ is Borel. The set

$$Z_n = \bigg\{(x_1, y_1, \dots, x_n, y_n) \in \Gamma^n : \sum_{i=1}^{n} \big[c(x_{i+1}, y_i) - c(x_i, y_i)\big] < 0\bigg\} \cap C_n$$

is analytic: in fact, being the sum of a Borel function and an Π^1_1-function, the function

$$\sum_{i=1}^{n} \big[c(x_{i+1}, y_i) - c(x_i, y_i)\big]$$

is a Π^1_1-function.

The fact that there is no cyclic perturbation λ of π which lowers the cost $\mathcal{I}$ means that for all n

$$\sup \big\{m(Z_n) : m \in \Pi(\pi, \dots, \pi)\big\} = 0,$$

otherwise the projected measure λ given by (B.6) satisfies

$$\int c\lambda = \int_{Z_n} \frac{1}{n} \sum_{i=1}^{n} \big[c(x_{i+1}, y_i) - c(x_i, y_i)\big] m(dx_1 dy_1 \dots dx_n dy_n) < 0$$

contradicting optimality of π, as $\pi + \lambda$ would be a transference plan with lower cost.

Theorem B.2 implies that there are π-negligible sets $N_{n,i} \subset [0,1]^2$, $i = 1, \dots, n$, such that

$$Z_n \subset \bigcup_{i=1}^{n} (P_{2i-1,2i})^{-1}(N_{n,i}).$$

The set $\Gamma \setminus \cup_{i=1}^n N_{n,i}$ satisfies then (B.15) for cycles of length $I \leq n$. The c-cyclically monotone set Γ proving the proposition is finally $\Gamma \setminus \cup_n \cup_{i=1}^n N_{n,i}$. □

B.4. **Examples.** We give some examples.

Example B.17. Here we show that there are maxima of the problem (B.9) which are not symmetric, and for which the projection on a single component does not exhibit any cyclic structure. Consider the following example (since the measures are atomic, we use a matrix notation):

$$\lambda = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & -1 & 0 \\ -1 & 1 & 0 & -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

It is easy to verify that the maximum in the problem (B.9) with $n = 2$ is 2, by just considering the functions

$$h_1 = h_3 = 1 - \chi_{\mathrm{supp}\lambda^+}, \quad h_2 = h_4 = 1 - \chi_{\mathrm{supp}\lambda^-} + \delta_{\{3,2\}},$$

and that a maximizer is the measure:

$$\bar{\pi}\llcorner_{\tilde{D}_2} = \delta_{(\{3,1\},\{3,2\},\{4,2\},\{4,1\})} + \delta_{(\{2,2\},\{2,5\},\{3,5\},\{3,2\})}.$$

It follows that $\lambda_2 \neq (P_1, P_2)_\sharp \bar{\pi}\llcorner_{\tilde{D}_2} - (P_3, P_4)_\sharp \bar{\pi}$: indeed

$$(P_1,P_2)_\sharp \bar{\pi}\llcorner_{\tilde{D}_2} - (P_3,P_4)_\sharp \bar{\pi}\llcorner_{\tilde{D}_2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Conversely the symmetrized measure yields

$$\lambda_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & -1/2 & 0 & 0 \\ 1/2 & -1 & 0 & 0 & 1/2 & 0 & 0 \\ -1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

This example proves also that we do not have uniqueness, by just observing that

$$\begin{aligned} \big\{\pi : \pi(D_2) = 2, \pi \in \Pi(\lambda^+, \lambda^-, \lambda^+, \lambda^-)\big\} \\ = \Big\{ & \alpha_1 \delta_{(\{3,1\},\{3,2\},\{4,2\},\{4,1\})} + \alpha_2 \delta_{(\{4,2\},\{4,1\},\{3,1\},\{3,2\})} \\ & + \alpha_3 \delta_{(\{2,2\},\{2,5\},\{3,5\},\{3,2\})} + \alpha_4 \delta_{(\{3,5\},\{3,2\},\{2,2\},\{2,5\})}, \alpha_i \geq 0, \sum_{i=1}^4 \alpha_i = 1 \Big\}. \end{aligned}$$

Hence the symmetrized set of π and the projected set are

$$\begin{aligned} \Big\{ & \alpha_1 \Big(\delta_{(\{3,1\},\{3,2\},\{4,2\},\{4,1\})} + \delta_{(\{4,2\},\{4,1\},\{3,1\},\{3,2\})} \Big) \\ & + \alpha_2 \Big(\delta_{(\{2,2\},\{2,5\},\{3,5\},\{3,2\})} + \delta_{(\{3,5\},\{3,2\},\{2,2\},\{2,5\})} \Big), \alpha_i \geq 0, \sum_{i=1}^2 \alpha_i = \frac{1}{2} \Big\}, \end{aligned}$$

$$\left\{ \alpha_1 \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \alpha_i \geq 0, \sum_{i=1}^2 \alpha_i = 1 \right\}.$$

Example B.18. Here we decompose the measure

$$\lambda := \begin{bmatrix} 0 & 0 & 0 & 0 & -m_1 & m_1 \\ -m_1 & 0 & 0 & m_1 & m_1 & -m_1 \\ 0 & 0 & 0 & -m_2 & m_1 & 0 \\ m_1 & -m_1 & 0 & 0 & 0 & 0 \\ 0 & m_1 & -m_1 & 0 & 0 & 0 \\ 0 & 0 & m_1 & 0 & -m_1 & 0 \end{bmatrix}$$

into an essentially cyclic and an acyclic part in two different ways, and the two acyclic part will not even have the same mass. Let

$$m_1 := (\mathbb{I}, \mathbb{I})_\sharp \mathcal{L}^1 \llcorner_{[0,a]}, \quad m_2 := (\mathbb{I} + \alpha \mod a, \mathbb{I})_\sharp \mathcal{L}^1 \llcorner_{[0,a]},$$

with $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Depending on $n = 2$ or $n = 4$ we obtain the following two decompositions:

$$\lambda = \begin{bmatrix} 0 & 0 & 0 & 0 & -m_1 & m_1 \\ 0 & 0 & 0 & 0 & m_1 & -m_1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -m_1 & 0 & 0 & m_1 & 0 & 0 \\ 0 & 0 & 0 & -m_2 & m_1 & 0 \\ m_1 & -m_1 & 0 & 0 & 0 & 0 \\ 0 & m_1 & -m_1 & 0 & 0 & 0 \\ 0 & 0 & m_1 & 0 & -m_1 & 0 \end{bmatrix},$$

$$\lambda = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -m_1 & 0 & 0 & 0 & m_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ m_1 & -m_1 & 0 & 0 & 0 & 0 \\ 0 & m_1 & -m_1 & 0 & 0 & 0 \\ 0 & 0 & m_1 & 0 & -m_1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & -m_1 & m_1 \\ 0 & 0 & 0 & m_1 & 0 & -m_1 \\ 0 & 0 & 0 & -m_2 & m_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The first measure is cyclic and the second is acyclic, because of m_2.

Appendix C. The c-cyclically monotone relation

Let c be a $\Pi^1_1([0,1]^2; [0,+\infty])$-function and let Γ be a c-cyclically monotone σ-compact set such that $c \llcorner_\Gamma$ is Borel and real valued. In the following this will be the set where a transference plan is concentrated.

The next definition is not the standard one, but it is useful for our construction.

Definition C.1 (Cyclically Monotone Envelope)**.** For a given function $f : [0,1] \to (-\infty, +\infty]$ define the *c-cyclically monotone envelope* of f as

(C.1)

$$\phi(x) = \begin{cases} \inf \left\{ \displaystyle\sum_{i=0}^{I} c(x_{i+1}, y_i) - c(x_i, y_i) + f(x_0), (x_i, y_i) \in \Gamma, x_{I+1} = x, I \in \mathbb{N} \right\} & \text{if the infimum is} < +\infty \\ -\infty & \text{otherwise} \end{cases}$$

Similarly, for a given function $g : [0,1] \to (-\infty, +\infty]$ define the c_{-1}*-cyclically monotone envelope* of g as

(C.2)

$$\psi(y) = \begin{cases} \inf \left\{ \displaystyle\sum_{i=0}^{I} c(x_i, y_{i+1}) - c(x_i, y_i) + g(y_0), (x_i, y_i) \in \Gamma, y_{I+1} = y, I \in \mathbb{N} \right\} & \text{if the infimum is} < +\infty \\ -\infty & \text{otherwise} \end{cases}$$

In the following we will denote them by

$$\mathtt{C}(f) \qquad \text{and} \qquad \mathtt{C}_{-1}(g).$$

Moreover, we will often call the first case of formulas (C.1), (C.2) as the inf-formula.

Lemma C.2. *If f, g belong to the Δ^1_n-pointclass with $n \geq 2$, then the functions $\phi, \psi : [0,1] \to [-\infty, +\infty)$ belong to the Δ^1_{n+1}-pointclass. Moreover $\phi(x) \leq f(x)$, $\psi(y) \leq g(y)$ for $x \in P_1(\Gamma)$, $y \in P_2(\Gamma)$.*

Proof. The second part of the lemma holds trivially, because of the particular path $(x_i, y_i) = (x, y) \in \Gamma$ for all i.

Consider the function

$$\phi_I(x_0, y_0, \dots, x_I, y_I, x) = \sum_{i=0}^{I} c(x_{i+1}, x_i) - c(x_i, y_i) + f(x_0), \quad (x_i, y_i) \in \Gamma,\ x_{I+1} = x.$$

Being the sum of the Π^1_1 functions $c(x_{i+1}, x_i) - c(x_i, y_i)$ ($c\llcorner\Gamma$ is Borel) with the Δ^1_n-function f, the function $\phi_I(x_0, y_0, \dots, x_I, y_I, x)$ is Δ^1_n with $n \geq 2$.

If $g(x, y)$ is a Δ^1_n-function, then $\tilde{g}(x) = \inf_y g(x, y)$ satisfies

$$\tilde{g}^{-1}(-\infty, s) = P_1(g^{-1}(-\infty, s)) \in \Sigma^1_n,$$

so that $\tilde{g}$ is in the Π^1_n-pointclass.

It follows that

$$\phi_I(x) = \inf\left\{ \sum_{i=0}^{I} c(x_{i+1}, x_i) - c(x_i, y_i) + f(x_0), (x_i, y_i) \in \Gamma, x_{I+1} = x \right\}$$

is Π^1_n, and finally $\inf_I \phi_I(x)$ is also Π^1_n. We conclude the proof by just observing that the set $\{x : \inf_I \phi_I(x) = +\infty\}$ is in Π^1_n, being the countable intersection of the Π^1_n-sets $\{x : \inf_I \phi_I(x) > k\}$. Hence $\{x : \inf_I \phi_I(x) < +\infty\} \in \Sigma^1_n$, so that the conclusion follows from the fact that $\Delta^1_{n+1} \supset \Sigma^1_n \cup \Pi^1_n$ and it is a σ-algebra. □

Remark C.3. In the case $n = 1$ the same proof shows that ϕ, ψ are $\mathcal{A}$-functions.

Definition C.4. A function $f : [0, 1] \to [-\infty, +\infty]$ is *c-cyclically monotone* if for all $x, x' \in [0, 1]$ such that $f(x) > -\infty$ and for all $(x_i, y_i) \in \Gamma$, $i = 0, \dots, I$, $x_0 = x$, $x_{I+1} = x'$, it holds

$$f(x') \leq f(x) + \sum_{i=0}^{I} c(x_{i+1}, y_i) - c(x_i, y_i).$$

Similarly, a function $g : [0, 1] \to [-\infty, +\infty]$ is *c_{-1}-cyclically monotone* if for all $y, y' \in [0, 1]$ such that $g(y) > -\infty$ and for all $(x_i, y_i) \in \Gamma$, $i = 0, \dots, I$, $y_0 = y$, $y_{I+1} = y'$, it holds

$$g(y') \leq g(y) + \sum_{i=0}^{I} c(x_i, y_{i+1}) - c(x_i, y_i).$$

The following are well known results: we give the proof for completeness. We recall that for any function $h : [0, 1] \mapsto [-\infty, +\infty]$ the set F_h is the set where h is finite:

$$F_h := h^{-1}(\mathbb{R}) = \big\{x \in [0, 1] : h(x) \in \mathbb{R}\big\}. \tag{C.3}$$

Lemma C.5. *Let $f : [0, 1] \to (-\infty, +\infty]$ ($g : [0, 1] \to (-\infty, +\infty]$). Then following holds:*

1. *The function $\phi := \mathtt{C}(f)$ ($\psi := \mathtt{C}_{-1}(g)$) defined in* (C.1) *(in* (C.2)*) is c-cyclically monotone (c_{-1}-cyclically monotone).*
2. *If f is c-cyclically monotone (g is c_{-1}-monotone), then $\phi(x) = f(x)$ on $F_f \cap P_1(\Gamma)$ ($\psi(x) = g(x)$ on $F_g \cap P_2(\Gamma)$).*
3. *If we define the function*

$$g'(y) = \begin{cases} c(x, y) - \phi(x) & (x, y) \in (F_\phi \times [0, 1]) \cap \Gamma \\ +\infty & \textit{otherwise} \end{cases} \quad \left(f'(x) = \begin{cases} c(x, y) - \psi(y) & (x, y) \in ([0, 1] \times F_\psi) \cap \Gamma \\ +\infty & \textit{otherwise} \end{cases} \right) \tag{C.4}$$

then g' is c_{-1}-cyclically monotone (f' is c-cyclically monotone) and belongs to the Δ^1_{n+1}-pointclass if f is in the Δ^1_n-pointclass (belongs to the Δ^1_n-pointclass if g is in the Δ^1_{n+1}-pointclass).

A part of the statement is that $c(x, y) - \phi(x)$ does not depend on x for fixed y in $(F_\phi \times [0, 1]) \cap \Gamma$ ($c(x, y) - \psi(y)$ does not depend on y for fixed x in $([0, 1] \times F_\phi) \cap \Gamma$).

Remark C.6. If ϕ, ψ are $\mathcal{A}$-functions, it is fairly easy to see that g', f' are $\mathcal{A}$-functions.

Proof. The proof will be given only for ϕ, the analysis for ψ being completely similar.

Point (1). The first part follows by the definition: for any axial path as in Definition C.4 we have

$$\begin{aligned}
&\phi(x) + \sum_{i=0}^{I} c(x_{i+1}, y_i) - c(x_i, y_i) \\
&= \inf \left\{ \sum_{i=0}^{I'} c(x_{i+1}, x_i) - c(x_i, y_i) + f(x_0), (x_i, y_i) \in \Gamma, x_{n+1} = x, I' \in \mathbb{N} \right\} + \sum_{i=0}^{I} c(x_{i+1}, y_i) - c(x_i, y_i) \\
&\geq \inf \left\{ \sum_{i=0}^{I'+I} c(x_{i+1}, x_i) - c(x_i, y_i) + f(x_0), (x_i, y_i) \in \Gamma, x_{I'+I+1} = x', I' \in \mathbb{N} \right\} \geq \phi(x').
\end{aligned}$$

Notice that we have used that $\phi(x) > -\infty$ to assure that its value is given by the inf-formula.

Point (2). The second point follows by the definition of c-cyclical monotonicity: first of all, if $x \in F_f \cap P_1(\Gamma)$, the value of ϕ is computed by the inf-formula in (C.1) by Lemma C.2. Then we have from the c-cyclical monotonicity of f

$$f(x) \leq \sum_{i=0}^{I} c(x_{i+1}, y_i) - c(x_i, y_i) + f(x_0), \quad x_0 \in F_f, (x_i, y_i) \in \Gamma, x_{I+1} = x.$$

Hence we obtain $\phi(x) \geq f(x)$, and using Lemma C.2 we conclude the proof of the second point.

Point (3). Assume that for y fixed there are $x, x' \in F_\phi$ such that $(x, y) \in \Gamma$ and

$$c(x, y) - \phi(x) \geq c(x', y) - \phi(x') + \epsilon.$$

Then, since $x, x' \in F_\phi$, there are points $(x_i, y_i) \in \Gamma$, $i = 0, \dots, I$, $x_{I+1} = x$ such that

$$\sum_{i=0}^{I} c(x_{i+1}, x_i) - c(x_i, y_i) + f(x_0) < \phi(x) + \frac{\epsilon}{2}.$$

Add then the point $(x_{I+1}, y_{I+1}) = (x, y) \in \Gamma$ to the previous path: the definition of ϕ implies then for $x_{I+2} = x'$

$$\begin{aligned}
\phi(x') \leq\ & \sum_{i=0}^{I+1} c(x_{i+1}, x_i) - c(x_i, y_i) + f(x_0) \\
&= c(x', y_{I+1}) - c(x_{I+1}, y_{I+1}) + \sum_{i=0}^{I} c(x_{i+1}, y_i) - c(x_i, y_i) + f(x_0) \\
&< c(x', y) - c(x, y) + \phi(x) + \frac{\epsilon}{2},
\end{aligned}$$

yielding a contradiction. This shows that the definition of g makes sense.

The proof of the c-cyclical monotonicity is similar: assume that there exist points $(x_i, y_i) \in \Gamma$, $i = 0, \dots, I$, such that $g'(y_0) > -\infty$ and

$$g'(y') > g'(y) + \sum_{i=0}^{I} c(x_i, y_{i+1}) - c(x_i, y_i), \quad y_0 = y,\ y_{I+1} = y'$$

Using the fact that $g'(y), g'(y') > -\infty$, it follows that there exists $(x, y), (x', y') \in (F_\phi \times [0, 1]) \cap \Gamma$ such that $g'(y) = c(x, y) - \phi(x)$, $g'(y') = c(x', y') - \phi(x')$ so that for $(x_{I+1}, y_{I+1}) = (x', y')$, $(x_0, y_0) = (x, y)$

$$\begin{aligned}
g'(y') >\ & g'(y) + \sum_{i=0}^{I} c(x_i, y_{i+1}) - c(x_i, y_i) \\
&= c(x, y) - \phi(x) + c(x', y') + \sum_{i=1}^{I+1} c(x_{i-1}, y_i) - c(x_i, y_i) - c(x_0, y_0) \\
&\geq c(x', y') - \phi(x) - \phi(x') + \phi(x) \\
&= c(x', y') - \phi(x') = g'(y'),
\end{aligned}$$

yielding a contradiction. We have used the c-cyclical monotonicity of ϕ.

Finally, since $c_{\llcorner\Gamma}$ is Borel, then it follows immediately that g' is in the Δ^1_{n+1}-pointclass. □

For fixed $(\bar x, \bar y) \in \Gamma$, we can thus define recursively for $i \in \mathbb{N}_0$ the following sequence of functions ψ_{2i}, ϕ_{2i+1}.

(1) Set $\psi_0(y; \bar x, \bar y) = -1\!\!1_{\bar y}(y) \in \Delta^1_1$.

(2) Assume that $\psi_{2i}(\bar x, \bar y) \in \Delta^1_{2i+1}$ is given. For $i \in \mathbb{N}_0$, define then the function $\phi_{2i+1}(x; \bar x, \bar y)$ as

$$\phi_{2i+1}(\bar x, \bar y) = \mathtt{C}\big((\psi_{2i})'(\bar x, \bar y)\big) \in \Delta^1_{2i+2}, \tag{C.5}$$

where $(\psi_{2i}(\bar x, \bar y))' \in \Delta^1_{2i+1}$ is defined in (C.4).

(3) Similarly, if $\phi_{2i+1}(\bar x, \bar y) \in \Delta^1_{2i+2}$ is given, define

$$\psi_{2i+2}(\bar x, \bar y) = \mathtt{C}_{-1}\big((\phi_{2i+1})'(\bar x, \bar y)\big) \in \Delta^1_{2i+3}. \tag{C.6}$$

Note that ϕ_{2i+1} is a Δ^1_{2i+2}-function, ψ_{2i+2} is a Δ^1_{2i+3}-function for $i \in \mathbb{N}_0$ (Lemma C.5), so that the sets

$$A_{2i+1}(\bar x, \bar y) = F_{\phi_{2i+1}(\bar x, \bar y)}, \quad B_{2i+2}(\bar x, \bar y) = F_{\psi_{2i+2}(\bar x, \bar y)}, \quad i \in \mathbb{N}_0, \tag{C.7}$$

are in Δ^1_{2i+2}, Δ^1_{2i+3}, respectively.

From Lemma C.5 it follows the next corollary.

Corollary C.7. *If $\phi_{2i+1}(x, \bar x, \bar y)$, $\psi_{2i}(y, \bar x, \bar y)$ are constructed by* (C.5), (C.6) *and $A_{2i+1}(\bar x, \bar y)$, $B_{2i}(\bar x, \bar y)$ are defined by* (C.7)*, then the following holds:*

(1) *$A_{2i+1} \subset A_{2j+1}$, $B_{2i} \subset B_{2j}$ if $i \leq j$, and*

$$\phi_{2j+1}(\bar x, \bar y)_{\llcorner A_{2i+1}(\bar x, \bar y)} = \phi_{2i+1}(\bar x, \bar y), \quad \psi_{2j}(\bar x, \bar y)_{\llcorner A_{2i}(\bar x, \bar y)} = \psi_{2i}(\bar x, \bar y).$$

(2) *$A_1(\bar x, \bar y) \supseteq P_1(\Gamma \cap ([0,1] \times \{\bar y\}))$ and in general*

$$A_{2i+1}(\bar x, \bar y) \supseteq P_1\Big(\big([0,1] \times B_{2i}(\bar x, \bar y)\big) \cap \Gamma\Big), \quad B_{2i+2}(\bar x, \bar y) \supseteq P_2\Big(\big(A_{2i+1}(\bar x, \bar y) \times [0,1]\big) \cap \Gamma\Big).$$

(3) *On the set $(A_{2i+1}(\bar x, \bar y) \times A_{2j}(\bar x, \bar y)) \cap \Gamma$ it holds*

$$\phi_{2i+1}(x, \bar x, \bar y) + \psi_{2j}(x, \bar x, \bar y) = c(x, y).$$

Proof. Point (1). Point (3) of Lemma C.5 implies that at each step we are applying formula (C.1) to the c-cyclically monotone function $c(x, y) - \psi_{2i}(y)$ or the c_{-1}-cyclically monotone $c(x, y) - \phi_{2i+1}(y)$. From Point (2) of the same lemma we deduce Point (1).

Point (2). The second point is again a consequence of the c-cyclically monotonicity or c_{-1}-cyclically monotonicity of the functions $c(x, y) - \psi_{2i}(y)$, $c(x, y) - \phi_{2i+1}(y)$ on the set $([0,1] \times B_{2i}(\bar x, \bar y)) \cap \Gamma$, $(A_{2i+1}(\bar x, \bar y) \times [0,1]) \cap \Gamma$, respectively.

Point (3). The last point follows from Point (2) by Lemma C.5. □

For all $(x, y) \in \Gamma$, define the set $\Gamma_{(x,y)}$ as

$$\Gamma_{(x,y)} := \Gamma \cap \bigg(\bigcup_i A_{2i+1}(x, y) \times B_{2i}(x, y)\bigg).$$

Observe that under (PD) $\Gamma_{(x,y)}$ is measurable for all Borel measures (Section B.1).

We then define the following relations in $[0,1]^2$.

Definition C.8 (c-cyclically monotone relation)**.** We say that $(x, y)R(x', y')$ if $(x', y') \in \Gamma_{(x,y)}$. We call this relation R the *c-cyclically monotone relation*.

Clearly $\bar E \subset R$, where $\bar E$ is the closed cycle equivalence relation given in Definition 5.4: actually the equivalence class of $(\bar x, \bar y)$ w.r.t. $\bar E$ is already contained in $(A_1 \times [0,1]) \cap \Gamma$.

Remark C.9. The following are easy observations.

(1) If $(x, y)R(x', y')$, then from Point (2) of Corollary C.7 also $(x, y)R(\Gamma \cap (\{x'\} \times [0,1]))$ and $(x, y)R(\Gamma \cap ([0,1] \times \{y'\}))$: this means that Γ satisfies the crosswise condition w.r.t. R (Definition 2.3). In particular to characterize R it is enough to define the projected relations

$$xR_1x' \;\Leftrightarrow\; x' \in \bigcup_{i \in \mathbb{N}_0} A_{2i+1}(x, y), \quad yR_2y' \;\Leftrightarrow\; y' \in \bigcup_{i \in \mathbb{N}_0} B_{2i}(x, y).$$

(2) The relation R is nor transitive neither symmetric, as the following example shows (see Figure 7).
Consider the cost

$$c(x,y) = \begin{cases} 0 & (x,y) \in A \\ 1 - \sqrt{x-y-7/8} & 7/8 \leq y + 7/8 \leq x \leq 1 \\ +\infty & \text{otherwise} \end{cases}$$

where

$$A = \Big\{(0,0),(0,1/4),(0,1/2),(1/4,1/2),(1/2,1/2),(1/2,3/4),(1/2,1),(3/4,1),(1,1)\Big\}.$$

Let Γ be the set

$$\Gamma = \big\{(0,1/4),(1/4,1/2),(1/2,3/4),(3/4,1)\big\} \cup \big\{(x,x-7/8), x \in [7/8,1]\big\}.$$

It is easy to see that

$$\begin{aligned} \Gamma_{(1/4,1/2)} &= \Gamma \cap (\{0,1/4,1/2,3/4,7/8\} \times [0,1]) \\ &\neq \Gamma \cap ([0,1] \times \{1/8,1/4,1/2,3/4,1\}) = \Gamma_{(1/2,3/4)}. \end{aligned}$$

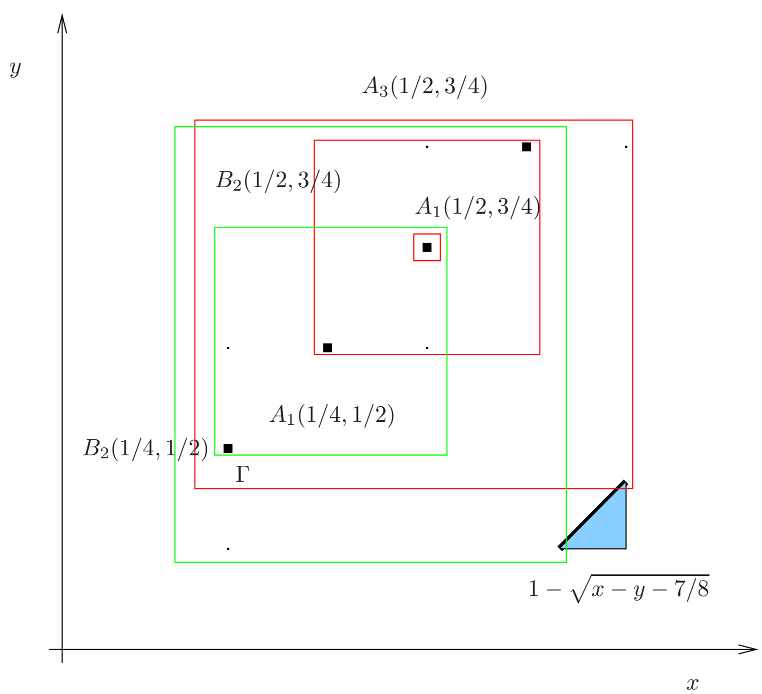


Figure 7: The cost of Point (2) of Remark C.9.

(3) Another possible definition can for example be the following symmetric relation on Γ.

Definition C.10. We say that $(x',y')R(x'',y'')$ if there exists Borel functions $\phi,\psi : [0,1] \mapsto \mathbb{R} \cup \{-\infty\}$ such that

$$\phi(x') + \psi(y') = c(x',y'), \quad \phi(x'') + \psi(y'') = c(x'',y''), \quad \phi(x) + \psi(y) \leq c(x,y)\,\forall (x,y).$$

However, the following points are in order.

(a) The relation R depends deeply on the choice of Γ.

(b) We observe that even if $R_x = \{y : yRx\} = [0,1]$ for some x, this does not mean that the measure is optimal. As an example, consider (Figure 8)

$$c(x,y) = \begin{cases} 1 & 0 < x = y < 1 \\ 0 & 1 > x = y - \alpha \mod 1 \\ 0 & y = 0 \\ +\infty & \text{otherwise} \end{cases}$$

with $\alpha \in [0,1] \setminus \mathbb{Q}$, and the transport problem $\mu = \delta_1 + \mathcal{L}^1$, $\nu = \delta_0 + \mathcal{L}^1$. The transference plan $\pi = \delta_{(1,0)} + (\mathbb{I}, \mathbb{I})_\sharp \mathcal{L}^1$ is clearly not optimal, but since the set

$$\Gamma = \big\{(x,x), x \in [0,1)\big\} \cup \{(1,0)\}$$

has not closed cycles, it follows that it is c-cyclically monotone and moreover $R_1 = [0,1]$.

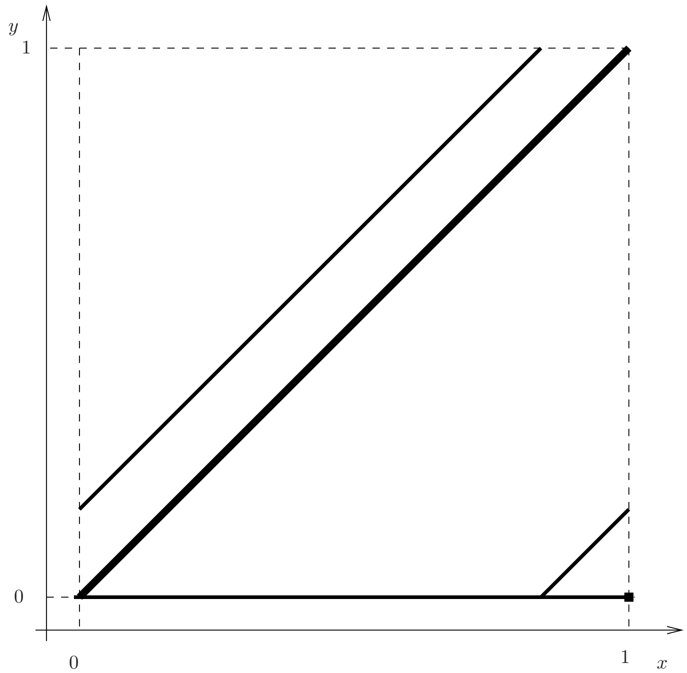


Figure 8: The cost function considered in Point (d) at Page 53

The main use of the c-cyclically monotone relation R is that any crosswise equivalence relation whose graph is contained in R and such that the disintegration is strongly consistent can be use to apply Theorem 5.6: the relation $\bar{E}$ of Definition 5.4 is a possible choice. Note that the strong consistency of the disintegration allows to replace the universally measurable equivalence classes with Borel one, up to a π-negligible set.

Remark C.11. Under (CH) we can give a procedure to construct an equivalence relation $E' \subset R$ maximal w.r.t. inclusion: if R_α, $\alpha \in \omega_1$, is an ordering of the partition $R_{(\bar{x},\bar{y})} = \{(x,y) : (\bar{x},\bar{y})R(x,y)\}$, one then defines the partition

$$E_\alpha = \Gamma \cap \bigg[\bigg(P_1 R_\alpha \setminus \bigcup_{\beta<\alpha} P_1 R_\beta\bigg) \times [0,1]\bigg].$$

Being R_β universally measurable and $\sharp\{\beta < \alpha\} = \omega_0$, we have that each E_α is universally measurable. Moreover it is a partition, and from the definition of R it follows that in each class there are optimal ϕ, ψ. Finally it is clearly maximal w.r.t. graph inclusion among all equivalence relations containing $\bar{E}$ and contained in R.

However, the graph of the above well ordering $\preccurlyeq$ of $\mathbb{R}$ cannot be $m \otimes m$-measurable for any non purely atomic measure m (see Example C.12 below).

Example C.12. Under (CH), there exists a well ordering $\preccurlyeq$ of $[0,1]$ such that $\sharp\{x : \ x \preccurlyeq y\} \leq \aleph_0$ (see the construction of Remark C.11). Denoting with m_a the absolutely continuous part of m, one has therefore

$$m_a(\{x : x \preccurlyeq y\}) = 0$$

for all $y \in \mathbb{R}$. If the graph R of $\preccurlyeq$ is $m \otimes m$-measurable, then it is $m_a \otimes m_a$-measurable and by Fubini Theorem $m_a \otimes m_a(R) = 0$ (see Lemma 4.14). This means that m is purely atomic.

Appendix D. Notation

$\mathcal{B}$ or $\mathcal{B}(X)$	Borel σ-algebra of the topological space $(X, \mathcal{T})$
$\mathcal{M}(X)$ or $\mathcal{M}(X, \Omega)$	signed measures on a measurable space (X, Ω)
$\mathcal{M}^+(X)$ or $\mathcal{M}^+(X, \Omega)$	positive measures on a measurable space (X, Ω)
$\mathcal{P}(X)$ or $\mathcal{P}(X, \Omega)$	probability measures on a measurable space (X, Ω)
$L(\mu; J)$	μ-measurable maps from the measure space (X, Ω, μ) to $J \subset \mathbb{R} \cup \{\pm\infty\}$
$\Pi(\mu_1, \dots, \mu_I)$	$\pi \in \mathcal{P}(\Pi_{i=1}^I X_i, \otimes_{i=1}^I \Sigma_i)$ with marginals $(P_i)_\sharp \pi = \mu_i \in \mathcal{P}(X_i)$
$\Pi^{\leq}(\mu_1, \dots, \mu_I)$	$\pi \in \mathcal{M}(\Pi_{i=1}^I X_i, \otimes_{i=1}^I \Sigma_i)$, $\pi \geq 0$, with $(P_i)_\sharp \pi \leq \mu_i \in \mathcal{P}(X_i)$
$\Pi^f(\mu, \nu)$	$\pi \in \Pi(\mu, \nu)$ for which $\mathcal{I}(\pi) \in \mathbb{R}$
$\Pi^{\mathrm{opt}}(\mu, \nu)$	$\pi \in \Pi(\mu, \nu)$ for which $\mathcal{I}(\pi)$ is minimal
$P_{i_1 \dots i_I}$	projection of $x \in \Pi_{k=1,\dots,K} X_k$ into its $(i_1, \dots, i_I)$ coordinates, keeping order
$d\mu_2/d\mu_1$	the Radon-Nikodym derivative of (the absolutely continuous part of) μ_2 w.r.t. μ_1
$h_\sharp \mu$	the push forward of the measure μ through h, $h_\sharp\mu(A) = \mu(h^{-1}(A))$
$\mathbf{P}(X)$	power set of X
$\mathcal{I}(\pi)$	cost function (1.2)
χ_A	the characteristic function of A, $x \mapsto \delta_x(A)$
$\mathbb{1}_A$	the indicator function of A, $\mathbb{1}_A(x) = \frac{1-\chi_A(x)}{\chi_A(x)} \in \{0, +\infty\}$
$A \vartriangle B$	the symmetric difference between two sets A, B
$\trianglelefteq$	the lexicographic ordering (4.3) on $[0,1]^\alpha$, α ordinal number
$\mathrm{dist}(A, B)$	distance defined in B.2
$\mathrm{graph}(f)$	the graph of the function $f : X \to Y$, $\mathrm{graph}(f) = \{(x, y), y = f(x)\} \subset X \times Y$
$\mathrm{epi}(f)$	epigraph of function f, $\mathrm{epi}(f) = \{(x, y), y \geq f(x)\} \subset X \times \mathbb{R}$
$\mathbb{I}$, $\mathbb{I}_d$	identity operator on a set and on the space $\mathbb{R}^d$
Λ	measures $\lambda \in \mathcal{M}([0,1]^d)$ with 0 marginals, see (B.4)
$\mathbb{N}, \mathbb{N}_0$, $\mathbb{Q}$, $\mathbb{R}$	natural numbers, natural numbers with 0, rational numbers, real numbers
Γ	c-cyclically monotone σ-compact subset of $[0,1]^2$
$\Gamma(A)$, $\Gamma^{-1}(B)$	the sets $\Gamma(A) = P_2(\Gamma \cap P_1^{-1}(A))$, $\Gamma(B) = P_1(\Gamma \cap P_2^{-1}(B))$
C_n	configuration set of n-cycles (B.4)
D_n	phase set of n-cycles (B.4)
q	projection operator (B.4)
$\tilde{D}_n$	reduced phase set of n-cycles (B.4)
T	cyclical permutation of coordinates, defined in Point (3) at Page 41
A_x, A^x	the sections $\{y : (x, y) \in A\}$, $\{y : (y, x) \in A\}$ for $A \subset X \times Y$
$f\llcorner_A$	the restriction of the function f to A
$\mu\llcorner_A$	the restriction of the measure μ to the σ-algebra $A \cap \Sigma$
$\mathcal{L}^d$	Lebesgue measure on $\mathbb{R}^d$
π^*	outer measure (B.1)
$\\|\mu\\|$	norm of $\mu \in \mathcal{M}([0,1])$
$\vert\mu\vert$, $(\mu)^+$	the nonnegative measures variation and positive part of $\mu \in \mathcal{M}(X)$
$\mu \wedge \nu$, $\mu \vee \nu$	the measures minimum and maximum of $\mu, \nu \in \mathcal{M}(X)$
Θ_π	π-completion of the Borel σ-algebra
$\Theta(\mu, \nu)$	$\Pi(\mu, \nu)$-universal σ-algebra (1.1)
xRy, R	a binary relation R over X
$\mathrm{graph}(R)$	graph of the binary relation R, $\mathrm{graph}(R) = \{(x, y) : xRy\} \subset X^2$

$x \sim y$ or xEy, E	an equivalence relation over X with graph E
$x^\bullet$	equivalence class of x, $x^\bullet = E_x$
$A^\bullet$	saturated set for an equivalence relation, $A^\bullet = \cup_{x \in A} x^\bullet$
$X^\bullet$, $X/\sim$	quotient space of an equivalence relation
Σ^1_1, $\Sigma^1_1(X)$	the pointclass of analytic subsets of Polish space X, i.e. projection of Borel sets
Π^1_1	the pointclass of coanalytic sets, i.e. complementary of Σ^1_1
Σ^1_n, Π^1_n	the pointclass of projections of Π^1_{n-1}-sets, its complementary
Δ^1_n	the ambiguous class $\Sigma^1_n \cap \Pi^1_n$
$\mathcal{A}$	σ-algebra generated by Σ^1_1
$\mathcal{A}$-function	$f : X \to \mathbb{R}$ such that $f^{-1}((t, +\infty])$ belongs to $\mathcal{A}$
F_h	the set where the function h is finite (C.3)
$\Sigma^1_n(\Pi^1_n,\Delta^1_n)$-function	$f : X \to \mathbb{R}$ such that $f^{-1}((t, +\infty]) \in \Sigma^1_n$ (Π^1_n, Δ^1_n)

References

[1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford, 2000.

[2] L. Ambrosio and A. Pratelli. Existence and stability results in the L^1 theory of optimal transportation. In Springer, editor, *Optimal Transportation and Applications*, volume 1813 of *Lecture notes in mathematics*, 2001.

[3] M. Beiglböck, M. Goldstern, G. Maresch, and W. Schachermayer. Optimal and better transport plans. *J. Funct. Anal.*, 256(6):1907–1927, 2009.

[4] M. Beiglböck and W. Schachermayer. Duality for borel measurable cost function. Preprint.

[5] S. Bianchini and F. Cavalletti. Monge problem in metric spaces. Preprint.

[6] L. Caravenna. Construction of a Borel c-envelope relation. Unpublished.

[7] R. G. Douglas. On extremal measures and subspace density. *Michigan Math. J.*, 11:243–246, 1964.

[8] H. Federer. *Geometric measure theory*. Berlin Springer-Verlag, 1969.

[9] D. H. Fremlin. *Measure theory*, volume 3. Torres Fremlin, Colchester, 2004. Measure algebras. Corrected second printing of the 2002 original.

[10] D. H. Fremlin. *Measure theory*, volume 4. Torres Fremlin, Colchester, 2006. Topological measure spaces. Part I, II. Corrected second printing of the 2003 original.

[11] K. Hestir and S. C. Williams. Supports of doubly stochastic measures. *Bernoulli*, 1(3):217–243, 1995.

[12] J. Hoffmann-Jorgensen. Existence of conditional probabilities. *Math. Scand.*, 28:257–264, 1971.

[13] H. G. Kellerer. Duality theorems for marginals problems. *Z. Wahrcsh. Verw. Gebiete*, 67(4):399–432, 1984.

[14] J. Lindenstrauss. A remark on doubly stochastic measures. *Amer. Math. Monthly*, 72:379–382, 1965.

[15] D. A. Martin and J. R. Steel. A proof of projective determinacy. *J. Amer. Math. Soc.*, 2(1):71–125, 1989.

[16] Y. N. Moschovakis. *Descriptive set theory*, volume 100 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 1980.

[17] A. Pratelli. On the sufficiency of c-cyclical monotonicity for optimality of transport plans. *Math. Z.*, 2007.

[18] S. T. Rachev and L. Rüschendorf. *Mass transportation problems. Vol. II*. Probability and its Applications (New York). Springer-Verlag, New York, 1998. Applications.

[19] S. M. Srivastava. *A course on Borel sets*, volume 180 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

SISSA, via Beirut 2, IT-34014 Trieste (ITALY)
E-mail address: bianchin@sissa.it

SISSA, via Beirut 2, IT-34014 Trieste (ITALY)
E-mail address: l.caravenna@sissa.it